

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2005, 13:04:08 ; Search time 42 Seconds
(without alignments)
1683.793 Million cell updates/sec

Title: US-10-783-417-2
Perfect score: 3869
Sequence: 1 MNQNDNNEVEIIDSHTSPY.....KLETIQKINTFFTNHTKTL 735

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: Pir1:*
- 2: Pir2:*
- 3: Pir3:*
- 4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1061.5	27.4	1180	2 A26858	parasporal crystal
2	1060.5	27.4	1180	2 I39870	parasporal crystal
3	893.5	23.1	934	2 B29838	parasporal crystal
4	758.5	19.6	1136	1 USB881	parasporal crystal
5	676	17.5	659	2 S10228	parasporal crystal
6	654	16.9	652	2 I39811	parasporal crystal
7	641.5	16.6	1138	2 A48944	parasporal crystal
8	618	16.0	1160	2 I40589	parasporal crystal
9	612	15.8	1157	1 S49247	parasporal crystal
10	592.5	15.3	652	2 A27323	parasporal crystal
11	576	14.9	1228	2 S00873	parasporal crystal
12	568.5	14.7	719	2 I39815	insecticidal prote
13	565.5	14.6	719	2 I39814	insecticidal prote
14	565.5	14.6	719	2 S25383	parasporal crystal
15	564	14.6	649	1 JH0261	parasporal crystal
16	553.5	14.3	719	2 I40590	crvV465 protein -
17	547.5	14.2	1174	2 S32649	parasporal crystal
18	544.5	14.1	1165	2 S11446	parasporal crystal
19	539.5	13.9	1178	1 USB8XH	parasporal crystal
20	528	13.6	1154	2 S39536	parasporal crystal
21	526	13.6	1177	2 A49785	parasporal crystal
22	516	13.3	1174	2 A42459	parasporal crystal
23	512.5	13.2	618	2 S11445	parasporal crystal
24	512.5	13.2	1171	2 A37829	parasporal crystal
25	512.5	13.2	1171	2 I40572	parasporal crystal
26	508.5	13.1	1176	2 A48970	parasporal crystal
27	505	13.1	1160	2 S32647	parasporal crystal
28	492.5	12.7	1166	2 S32645	parasporal crystal
29	487	12.6	655	2 JC7140	protoxin - Bacillu

30	483	12.5	1172	2 S32689	parasporal crystal
31	481	12.4	1181	2 A41052	parasporal crystal
32	480	12.4	823	2 S04181	parasporal crystal
33	476.5	12.3	934	2 A22798	parasporal crystal
34	475.5	12.3	1176	2 JT0241	parasporal crystal
35	475	12.3	1189	2 S00944	parasporal crystal
36	472	12.2	613	2 JG6033	mosquitocidal prot
37	471	12.2	1156	2 S19306	parasporal crystal
38	469.5	12.1	1176	2 JG2219	parasporal crystal
39	469	12.1	1155	2 I39838	parasporal crystal
40	465	12.0	1155	2 A26513	parasporal crystal
41	465	12.0	1155	2 JD0002	parasporal crystal
42	465	12.0	1156	2 A29125	parasporal crystal
43	464.5	12.0	1176	2 A22617	parasporal crystal
44	464.5	12.0	1176	2 S02215	parasporal crystal
45	464	12.0	1155	2 S02134	parasporal crystal

ALIGNMENTS

RESULT 1

A26858

parasporal crystal protein cry4Aa1 - Bacillus thuringiensis subsp. israelensis
N:Alternate names: Parasporal crystal protein cryIVA

C:Species: Bacillus thuringiensis subsp. israelensis

C:Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 05-Oct-2004

C:Accession: A26858; S48691

R:Ward, E.S.; Ellar, D.J.

Nucleic Acids Res 15, 7195, 1987

A:Title: Nucleotide sequence of a Bacillus thuringiensis var. israelensis gene encoding

A:Reference number: A26858; MUID:88015571; PMID:2821500

A:Accession: A26858

A:Molecule type: DNA

A:Residues: 1-1180 <WAR>

A:Cross-references: UNIPROT:P16480; UNIPARC:UPI000016EA42; GB:Y00423; NID:G40351; PIDN:CJ

A:Note: the authors translated the codon GCA for residue 308 as Thr

R:Nishimoto, T.; Yoshisue, H.; Ihara, K.; Sakai, H.; Komano, T.

FEBS Lett. 348, 249-254, 1994

A:Title: Functional analysis of block 5, one of the highly conserved amino acid sequences

A:Reference number: S48691; MUID:94307434; PMID:7913448

A:Accession: S48691

A:Molecule type: DNA

A:Residues: 667-676 <NIS>

A:Cross-references: UNIPARC:UPI000017819A

C:Superfamily: Parasporal crystal protein

C:Keywords: delta-endotoxin

Query Match 27.4%; Score 1061.5; DB 2; Length 1180;

Best Local Similarity 35.1%; Pred. No. 7e-62;

Matches 271; Conservative 118; Mismatches 278; Indels 105; Gaps 24;

Qy	1	MNQNDNNEVEIIDSHTSPYFERNNSDRPYTNPNQPLQNTNYKEWLNCOGNTQYG	60
Db	1	MNPYONKNEYETLASQKKL--NISNNTRYPIENSPKQLQSTNYKDWLNWCQQQYQYG	58
Qy	61	DNFTFASADTAAVSAGTIVSGTLLAGTGLTSGPIGIIIGAIISFGTLITVFPAG	120
Db	59	GDFTFIDS---GELSAVTIVGVTLTGFGFT----PLGL---ALIGETLIPVLFPAG	108
Qy	121	EQDKTVMQTFKMGEIFVDTPLTESIKQLKLTLEGFRQILQSYNTALDDWRKELQAP	180
Db	109	DQSNL--WSDFTQTKNIKKIEASTIYSNANKILNRSFNVISTYHNHLKTWE-----NNP	162
Qy	181	GLPSSALQQAALTUKIRFENVHNDIREIP--GFQLEYTKTLLPIYQAANFHLNLO	238
Db	163	NPQNTQDVRTQIQLVHVFQNVIPELVNSCPNPSPDCDYINILVLSYQAANLHLTVLN	222
Qy	239	QGAELADENWADIHPSQIEPNAGTSDDYVKLLKENIPKYSNYCANTYRTKLNLRDEP--	296
Db	223	QAVKFEAYLKNRQFDYLEP-LPTAIDYIPVLTKAIEDYTNVCVITYKKGLNLIKTTPOS	281
Qy	297	----NMKWSIFNDYRRYMTITVLDITISQFSLYDIKRYRDSIGGIEVKGIKNELTREIYTT	352

```
Db      282 NLDGNINWNTYRTKMTTAVLDVVALFPNYDVGY-----PIGVQSELTRREIYQV 333
QY      353 EINFDRLPOLRVQPNLATMEYNLTRASFLEQFIY-----TENTNFGNR-----401
Db      334 -LNFESPYKY--DFQYQEDSLTRRP-HLFTWLDLSLNFYEKAQTPNNFFTSYNNMFY 389
QY      402 -LVGISNRDAPYTSNITETLYGERGTSPKTKIRPFESYKVSIVTDQRQSPVPSPQPHF 460
Db      390 TLDNISQKSVFNGHNVTDKL--KSLGLATNIYI-----F 422
QY      461 IINOIEL---YLANGSSNNT-----LKYSAGSGLSNYO-NTTFFQPRKKDC 502
Db      423 LNVISLDNKLNDYNNISKMDFFITNGTRLLKEKELTAGSGQITVDNKNIFGLPILKRR 482
QY      503 NLVIDGCSFNFNNYSHILSHFSLFYYSYVIGLQQLDGLTGVLGWTHSSVDYRNAISDKI 562
Db      483 ENQGNPTLFTPDYNYSHILSFIKSLSPATYKTVY-----TPAWTHSSVDPKNTIYTHL 537
QY      563 ITWIPAIKGNLDNTNSKVIIEGPGHGTGNLVYLOSQRLEITCTETPNSTQSYFIRLRYATN 622
Db      538 TTQIPAVKANSLGTASKVQVQPGHGTGDLI--DFKDHFKITCQHSNFQOSYFIRIRYASN 595
QY      623 GAGNTLPNLSLITPGVIGIPQRLNNTFSGTNNYNNLQYGDYGFYQFPSTVTLPLNRPNIPF 682
Db      596 GSANTRAVINLSIPGVAEL--GMALNPTFSGTDYTNLKYDFQYLFESNEVFPAPNQNISL 654
QY      683 IFNRADV--SNSIILIDKIEFIPITSSMHQNRKQKLETIQTINKTINTFTNHTK 733
Db      655 VFNRSDVYNTTTLIDKIEFLPITRSIREDEKQKLETVQOIINTFYANPIK 706

RESULT 2
I39870
parasporal crystal protein ISRH4 - Bacillus thuringiensis
C:Species: Bacillus thuringiensis
C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 05-Oct-2004
C:Accession: I39870
R:Sen, K.; Honda, G.; Koyama, N.; Nishida, M.; Neki, A.; Sakai, H.; Himeno, M.; Kom. no,
Agric. Biol. Chem. 52, 873-878, 1988
A:Title: Cloning and nucleotide sequences of the two 130 kDa insecticidal protein genes
A:Reference number: I39869
A:Accession: I39870
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1180 <RES>
A:Cross-references: UNIPROT:P16480; UNIPARC:UPI0000001B1F; GB:D00248; NID:g216289; |IDN:
C:Superfamily: Parasporal crystal protein
C:Keywords: delta-endotoxin

Query Match 27.4%; Score 1060.5; DB 2; Length 1180;
Best Local Similarity 35.1%; Pred. No. 81e-62;
Matches 271; Conservative 118; Mismatches 278; Indels 105; Gaps 24;

QY      1 MNQNNNDNNEYIIDSHTSPYFPNRSNDRSPYTNPNQPLQNTNYKEWLNMCQGNTOYG 60
Db      1 MNPYQNKNEYETINASQKGL--NISNNYTRYPIENSPKQLQSTNYKDWLNMCQGNTOYG 58
QY      61 DNFEFPASADTTAAVSAGTIVSGTLGAGLTSISGPIGIIIGAILISFGTLITVFWPAG 120
Db      59 GDFETPIDS--GELSAITYIVGTVLTGFGFTT-----PLGL---ALIGFGTLIPVLPFAQ 108
QY      121 EODKTVMTQFIKMGEIFVDTPLTESIKQLQTLGFRQILQSYNTALDDWRKLRLOAP 180
Db      109 DOSNT--WSDFITQTKNIKKKIASTVISNANKILNRSFNVISTYHNLKTWE-----NNP 162
QY      181 GLPPSSALQQAALTALKIRFENVHNDPIREP--GFOLEYTKTLILLPIYAQAANFHLNLLQ 238
Db      163 NPQNTQDVQTQLVHYHFQNVIPELVNSCPNPNPSCDYNNILVLSYAQAANHLTLVNL 222
QY      239 QGAELADEWNADHPSQIEPNAGTSDDYKLLKKNIPKYSNYCANTYRTGLKNLRDEP-- 296
Db      223 QAVKFEAYLKNRQFDYLEP-LPTAIDYVPVLTKAIEDYNTNYCVTVYKKGLNLIKTPDS 281
```

```
QY      297 ----NMKWSIPNDYRRYMTTIVLDTISQFSLYDIKRYRDSIGGIEVKGIKNELTREIYTT 352
Db      282 NLDGNINWNTYRTKMTTAVLDVVALFPNYDVGY-----PIGVQSELTRREIYQV 333
QY      353 EINFDRLPOLRVQPNLATMEYNLTRASFLEQFIY-----TENTNFGNR-----401
Db      334 -LNFESPYKY--DFQYQEDSLTRRP-HLFTWLDLSLNFYEKAQTPNNFFTSYNNMFY 389
QY      402 -LVGISNRDAPYTSNITETLYGERGTSPKTKIRPFESYKVSIVTDQRQSPVPSPQPHF 460
Db      390 TLDNISQKSVFNGHNVTDKL--KSLGLATNIYI-----F 422
QY      461 IINOIEL---YLANGSSNNT-----LKYSAGSGLSNYO-NTTFFQPRKKDC 502
Db      423 LNVISLDNKLNDYNNISKMDFFITNGTRLLKEKELTAGSGQITVDNKNIFGLPILKRR 482
QY      503 NLVIDGCSFNFNNYSHILSHFSLFYYSYVIGLQQLDGLTGVLGWTHSSVDYRNAISDKI 562
Db      483 ENQGNPTLFTPDYNYSHILSFIKSLSPATYKTVY-----TPAWTHSSVDPKNTIYTHL 537
QY      563 ITWIPAIKGNLDNTNSKVIIEGPGHGTGNLVYLOSQRLEITCTETPNSTQSYFIRLRYATN 622
Db      538 TTQIPAVKANSLGTASKVQVQPGHGTGDLI--DFKDHFKITCQHSNFQOSYFIRIRYASN 595
QY      623 GAGNTLPNLSLITPGVIGIPQRLNNTFSGTNNYNNLQYGDYGFYQFPSTVTLPLNRPNIPF 682
Db      596 GSANTRAVINLSIPGVAEL--GMALNPTFSGTDYTNLKYDFQYLFESNEVFPAPNQNISL 654
QY      683 IFNRADV--SNSIILIDKIEFIPITSSMHQNRKQKLETIQTINKTINTFTNHTK 733
Db      655 VFNRSDVYNTTTLIDKIEFLPITRSIREDEKQKLETVQOIINTFYANPIK 706

RESULT 3
B29838
parasporal crystal protein - Bacillus thuringiensis subsp. israelensis (fragment)
C:Species: Bacillus thuringiensis subsp. israelensis
C:Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 05-Oct-2004
C:Accession: B29838
R:Thorne, L.; Garduno, F.; Thompson, T.; Decker, D.; Zounes, M.; Wild, M.; Walfield, A.M.
J. Bacteriol. 166, 801-811, 1986
A:Title: Structural similarity between the Lepidoptera- and Diptera-specific insecticidal
A:Reference number: A94672; MUID:86223796; PMID:3011746
A:Accession: B29838
A:Molecule type: DNA
A:Residues: 1-934 <THO>
A:Cross-references: UNIPROT:P09662; UNIPARC:UPI000001781A0; GB:M12662; NID:g143228

Query Match 23.1%; Score 893.5; DB 2; Length 934;
Best Local Similarity 31.9%; Pred. No. 6.3e-51;
Matches 245; Conservative 127; Mismatches 267; Indels 129; Gaps 30;

QY      1 MNQNNNDNNEYIIDSHTSPYFPNRSNDRSPYTNPNQPLQNTNYKEWLNMCQGNTOYG 60
Db      1 MNPYQNKNEYETINAFSNGF--SKSNNSRYPLANKENQPLKNTNYKDWLNVCQDNQOYG 58
QY      61 DNFEFPASADTTAAVSAGTIVSGTLGAGLTSISGPIGIIIGAILISFGTLITVFWPAG 120
Db      59 NNAGNFASSETIVGVSAGIIVGVTMLG-----AFAAP--VLAAGIISFGTLPIFW--QG 109
QY      121 EODKTVMTQFIKMGEIFVDTPLTESIKQLQTLGFRQILQSYNTALDDWRKLRKL 177
Db      110 SDPANVQDNLING----GRPIQEKNIINVLTSIVTPIKNQLDKQEFDFKKEPAR-- 163
QY      178 QAPGLPSSALQQAALTALKIRFENVHN---DFIREIFGFOLEYTKTLILLPIYAQAANFHL 234
Db      164 -----THANAKAVHDLFTTLEPIIDKLDMLKNNASYRIPT-----LPAYAQIATWHL 211
QY      235 NLLQGAELADEWNADHPSQIEPNAGTSDDYKLLKKNIPKYSNYCANTYRTGLKNLR 293
Db      212 NLLKHAATYNNIW---LQNOGINPSTFNSNYYGYLKRKIQETDYCIOTYNAGLTWIR 268
```

QY 294 DEPNKWSIPNDVRRYYMTITVLDTISFSLDYDIKRYRDSIGIEVKGIKNEIUREITYTTE 353

D6 269 TTNATNMYNTYRLMELTLVLDLIAFPNPVDPEKY-----PIGVKSILEIREVY-TN 319

QY 354 INFDRLPQLRVQNLAIMEYNLTRASPKLSFLBQFIETYENTN-----FGNRL 402

D6 320 VNSDTP-----RTITELENGLTR-NPTLFTHWNOGRFYTRNSRDILDPDYDFSFTGNQM 372

QY 403 VGISNRDAPTYNSNTITETLYGERTG---SPTTKTIIRPFESY----KVSIVTDQRSPPVSP 455

D6 373 -----APTHWDNRNI IWGAHVHGNIIISQDTSKVFPYVRNKPIDKVEIVRHREYSDI-- 424

QY 456 IQPHFIINQIELYLNGSSNTLKYSAGSL-SNTQNTFFOFPRKKOCLNLVDPGCSPNF 514

D6 425 -----YEMIFFNSSSEVFYSSSNSTENNYKRTDSYMI PKQTWN----- 464

QY 515 NNYSHILSFSLPYSYVIGLOILOLTGLVGTWTHSSVDYRNAISDKLIITMIPAIGNNL 574

D6 465 EYGHLLSYIKTDNYIFSVVRERRV---AFSWHTISVDFONTIDLDNIQTIALKALKV 521

QY 575 DTNSKVIEGFGHTGMLVYLQSQRLEBITCETPNSTOSFYIRLRVATNAGAGTLPNISLT 634

D6 522 SSDSKIVKGFGHTGGDLVILKSDMDPRVR-LKNVSRQYQVRI RYATNA-----PKTYTF 575

QY 635 IPGVGIPQRLNNTSGTNN--NLQYGDFGYQFPSTV-----TLP LNRPFI 683

D6 576 LTGIDTISVE-LPSTTSRQPNPATDLTYADFGYVTFPRTPVNKTPEGEDTLTMT-----L 629

QY 684 FNRADVSNILIADKIEFIPITSMHQNRKOKLETQTCKINTPTNH 731

D6 630 YGTPNHSYNI-YIDKIEFIPITQSULDYTEKONIEKTQKI VINDLFWNN 676

RESULT 4
USBS91

Parasporal crystal protein cry4Ba1 [validated] - Bacillus thuringiensis subsp. israelensis
N:Alternate names: delta-endotoxin Bt8; insecticidal protein ISRH3; mosquitocidal 130K P18
C:Species: Bacillus thuringiensis subsp. israelensis
C>Date: 30-Jun-1991 #sequence revision 30-Jun-1991 #text change 05-Oct-2004
A:Accession: S00398; A37587; J70315; A28541; I39869; I40584
R:Chungjatapornchai, W.; Hoefle, H.; Seurinck, J.; Anguthanasombat, C.; Vaeck, M.
Eur. J. Biochem. 173, 9-16, 1988
A>Title: Common features of Bacillus thuringiensis toxins specific for Diptera and Lepidoptera
A:Reference number: S00398; UID:88185334; PMID:2833395
A:Accession: S00398
A:Molecule type: DNA
A:Residues: 1-1136 <CH>
A:Cross-references: UNIPROT:P05519; UNIPARC:UIP0000001BI1d; EMBL:X07423; NID:g40353; PIDN:R.Yamamoto, T.; Watkinson, I.A.; Kim, L.; Sage, M.V.; Stratton, R.; Akande, N.; Li, Y.; Gene 66, 107-120, 1988
A>Title: Nucleotide sequence of the gene coding for a 130-kDa mosquitocidal protein of E. coli
A:Reference number: J70315; UID:88329719; PMID:2901387
A:Accession: A37587
A:Molecule type: DNA
A:Residues: 1-50, 'D', '52-64', 'S', '66-192', 'PHKCTRMVY', '202-204', 'C', '207-363', 'LVQIYLKFN', '374', 'YKFLEEQLVNYV', '573-593', 'N', '595-686', 'GIIS', '691-720', 'R', '722-822', 'LIIVSVRCA', '833-835', 'WD'
A:Cross-references: UNIPARC:UIP000016EA34; GB:M20242; NID:g142737; PIDN:AAA22337.1; PID:R.Yamamoto, T.; Watkinson, I.A.; Kim, L.; Sage, M.V.; Stratton, R.; Akande, N.; Li, Y.; Gene 66, 107-120, 1988
A>Title: The complete nucleotide sequence of a 130 kDa mosquitocidal larvicidal delta-endotoxin
A:Reference number: A28541; UID:88157738; PMID:2831510
A:Accession: A28541
A:Molecule type: DNA
A:Residues: 1-204, 'R', '206-1136 <TUN>
A:Cross-references: UNIPARC:UIP000016EA3C; EMBL:X07082; NID:g40309; PIDN:CAA30114.1; PID:R.Sen, K.; Honda, G.; Koyama, N.; Nishida, M.; Neki, A.; Sakai, H.; Himeno, M.; Komano, Agric. Biol. Chem. 52, 873-878, 1988
A>Title: Cloning and nucleotide sequences of the two 130 kDa insecticidal protein genes

[illegible]

Db 637 VLDETNQNLSEREVVNALFTNDK 662

RESULT 5
S10228
parasporal crystal protein cry3Ba1 - Bacillus thuringiensis (fragment)
N:Alternate names: coleopteran-active parasporal crystal protein; delta-endotoxin
C:Species: Bacillus thuringiensis
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 05-Oct-2004
C:Accession: S10228
R:Sick, A.; Gaertner, F.; Wong, A.
Nucleic Acids Res. 18, 1305, 1990
A:Title: Nucleotide sequence of a coleopteran-active toxin gene from a new isolate (f Ba
A:Reference number: S10228; MUID:90206811; PMID:2320431
A:Accession: S10228
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-659 <SIC>
A:Cross-references: UNIPROT:P17969; UNIPARC:UPI0000126C13; EMBL:X17123; NID:g40258; PIDN
C:Genetics:
A:Gene: cryIIIB
C:Superfamily: Parasporal crystal protein
C:Keywords: delta-endotoxin; toxin

Query Match 17.5%; Score 676; DB 2; Length 659;
Best Local Similarity 28.0%; Pred. No. 8.9e-37;
Matches 208; Conservative 116; Mismatches 286; Indels 134; Gaps 28;

QY 1 MNQNDNNEYEIIDSTSPYFNNRNSDRYPYTNPNQPLQNTNYKWLKMQGNTQYG 60
DB 9 MNPNN-RSEYDTIKVTPNSGLP---TNHQYPLADPNSTLELNKYEFLRMADNST-- 62
QY 61 DNFETPASADTTAAVAGATVSGTLGAGLGLTSISGPIGIIAIIISFGTLITVFWPAG 120
DB 63 ----EVLDSSTVKDAGTGVSVGQILGVWG-----VFPAGALTsfyQSFNLAIWPS- 110
QY 121 EQDKTWTQFIKMGEIFVDTPLTSTESIKQLKLTLEGPRQLQSYNTALDWRKLRQAP 180
DB 111 --DADPWKAFMAQVEVLIDKKIEYAKSKALAEQLQNNFEDYVNALDSWK----- 161
QY 181 GLPPSALQQAALTAKIRENVHNDFIREIPGQLETYKTLPIYAAANFHLNLQOG 240
DB 162 -APVNLRSRSDRIRELFSQAESHFNMPSPFAVSKFEVLFLPTVYAAANTHLLIKDA 220
QY 241 AELADEWNADIHPSQIEPNAGTSD--YKLLKENIPKYSNCANTYRTGLKLRDEPN 297
DB 221 QVFGGEWY-----SSEDIAEFYQRLKLTQQYTDHCNVNWNVGLNSLRGSTY 268
QY 298 MKWSIFNDYRRVMTITVLDTISQFSLYDIKRYRDSIGGIEVKGIKNELTREIYTTTEINF 357
DB 269 DAWVKENFRREMTLVLDLIVLFPFVDRLYS-----KGVTLETRDIFTDPI-FT 319
QY 358 RLPLQVQPNLAPMEYNLFRASPKLFSLEQFIPTNTNTNFGNRL-VGISNRDAPY--- 413
DB 320 LNALQVGYGTFSSIENSIRKP--HLFDYLRGIEFHT-----RLRPYSGKDSFNWWSG 370
QY 414 -----SNTIETLYGERTGSPFTTKTIRPFES---YKYSIVTDROSPPVSPIQPH 459
DB 371 NYVETRPSIGSNDTITSPFYGDKSIEPIQKL--SFDQKQVVRTIANTDIAAFPDKI--Y 426
QY 460 FIINQIEL-VLNGSSNNLTLYSAGGSLSNYQNTTFFQPRKDCNLVIDPGCS--PNFNN 516
DB 427 FGVTKVDFQYDDQKNET-----STQTYDSKRYNGYLGAQDSIDQLPEITDPELEKA 479
QY 517 YSHILSHFSLFTYSYVIGLQQLDITGVLGWTHSSVDRYNAISDKIITMTIPAKGNLND 576
DB 480 YSHQLNVAECFL-----MQDRGRTIPFTTWTHTSRVDFDFTIDAEEKITQLPVVKAYALS 533
QY 577 NKSIVBGPCHTGNLVYLO-----SQGRLEITCTPNSTQSYFRLRYATNGAGNTLPNIS 632
DB 534 GASIIEGPGFTGGNLLFLKSSNSIAKFKVTLNSAALLQRYRIRYAST-----TNLR 587
QY 633 LITPGVIGIPPPQRLNNTFSGTNNYLNQY--GDFGY--FQFPSTVTLPLNRNPIFENRAD 688

Db 588 LFV-----QSNNDFLVIYINKTNIDGDLTYQTDFEATS-----NSNMGSGDNTD 634

QY 689 -----VSNSTILIDKIERFIPI 704
DB 635 FIIGAESFVSNKIIYIDKIEFIPV 658

RESULT 6
I39811
parasporal crystal protein cry3Bb1 - Bacillus thuringiensis
N:Alternate names: Parasporal crystal protein cryIIIB2
C:Species: Bacillus thuringiensis
C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 05-Oct-2004
C:Accession: I39811
R:Donovan, W.P.; Ruper, M.J.; Slaney, A.C.; Malvar, T.; Gawron-Burks, M.C.; Johnson, T.B.
Appl. Environ. Microbiol. 58, 3921-3927, 1992
A:Title: Characterization of two genes encoding Bacillus thuringiensis insecticidal cryst
A:Reference number: I39811; MUID:93119147; PMID:1476436
A:Accession: I39811
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-652 <RES>
A:Cross-references: UNIPROT:Q06117; UNIPARC:UPI0000126C14; GB:M89794; NID:g142729; PIDN:I
C:Genetics:
A:Gene: cryIIIB2
C:Superfamily: Parasporal crystal protein

Query Match 16.9%; Score 654; DB 2; Length 652;
Best Local Similarity 27.8%; Pred. No. 2.5e-35;
Matches 206; Conservative 114; Mismatches 294; Indels 126; Gaps 27;

QY 1 MNQNDNNEYEIIDSTSPYFNNR--NSNDSRYPYTNPNQPLQNTNYKWLKMQC-QNT 57
DB 1 MNPNN-RSEHD-----TIKVTNSELQTNHNQYPLADPNSTLELNKYEFLRMTEDSST 54
QY 58 QYDGNFETPASADTTA-AVSGATVSGTLGAGLGLTSISGPIGIIAIIISFGTLITVF 116
DB 55 EVLDN-----STVKDAGTGVSVGQILGVWG-----VFPAGALTsfyQSFNLTI 99
QY 117 WPAGEQDKTWTQFIKMGEIFVDTPLTSTESIKQLKLTLEGPRQLQSYNTALDWRKLR 176
DB 100 WPS---DADPWKAFMAQVEVLIDKKIEYAKSKALAEQLQNNFEDYVNALSNWKK--- 153
QY 177 LQAPGPPSALQQAALTAKIRENVHNDFIREIPGQLETYKTLPIYAAANFHLNL 236
DB 154 -----TPLSRKRSQDRIRELFSQAESHFNMPSPFAVSKFEVLFLPTVYAAANTHLL 208
QY 237 LQGAELADEWNADIHPSQIEPNAGTSDDYKLLKENIPKYSNCANTYRTGLKLRDEP 296
DB 209 LKDAQVGEWGSY-----SEDVAEFYHRQLKLT-QQYTDHCNVNWNVGLNLRGST 259
QY 297 NMKWSIFNDYRRVMTITVLDTISQFSLYDIKRYRDSIGGIEVKGIKNELTREIYTTTEINF 356
DB 260 YDAWKVFNRRREMTLVLDLIVLFPFVDRLYS-----KGVTLETRDIFTDPI-F 310
QY 357 DRPLQVQPNLAPMEYNLFRASPKLFSLEQFIPTNTNTNFGNRL-VGISNRDAPY--- 405
DB 311 SLMTLOEYGFTELSIENSIRKP--HLFDYLRGIEFHTRLQPGYFGKDSFNWWSGNY--- 365
QY 406 SNRDAPYSNNTIETLYGERTGSPFTTKTIRPFES---YKYSIVTDROSPP-----VSP 455
DB 366 ETRPSIGSSKTIITSPFYGDKSTPEVQKL--SFDQKQVVRTIANTDIAAWPNGKVLYGVTK 423
QY 456 IQPHFIINQIELYNGSSNNLTLYSAGGSLSNYQNTTFFQPRKDCNLVIDPGCSNPN 515
DB 424 VD-----PSQYDDQKNETSTQTYDSKRYNGYVSAQDSIDQLPPEITD-----EPEL 470
QY 516 NYSHILSHFSLFTYSYVIGLQQLDITGVLGWTHSSVDRYNAISDKIITMTIPAKGNLND 575
DB 471 AYSHQLNVAECFL-----MQDRGRTIPFTTWTHTSRVDFDFTIDAEEKITQLPVVKAYALS 524
QY 576 TNSKIVBGPCHTGNLVYLO-----SQGRLEITCTPNSTQSYFRLRYATNGAGNTLPNI 631

Query Match	16.0%; Score 618; DB 2; Length 1160;
Best Local Similarity	21.46%; Pred. No. 1.4e-32;
Matches	211; Conservative 116; Mismatches 293; Indels 174; Gaps 35;
QY	5 NDNVEYELIDSHSTGYPYPRNRNSDGRYPVTNNPNQPLQNTNYKEWLNMCQG-NTOYGDNF 63
DB	4 NNQNEYELIIDA-LSP--TSVSDNSIRYPLANDQTWLNQNNYKDYLNKWTSTNAELSRNP 60
QY	64 ETFASADTIAAVSAGTVISGTLLAGIGGLTISGPIGIICAILISFGTLITVFWPAGEQD 123
DB	61 GTFISAQD--AVGTDIVSTIISGLG-----IPVLGEVPSILGSLGILLWPSNNEN 110
QY	124 KTVMTQFIKMGEIFVDTPLTFSIKQLKLTLSGFRQILQSNTALDDWRKLKRLQAQGLP 183
DB	111 --VWQIFMNRVEELIDQKILDSVRSRAADLANSRITAVEYQNALEDWRK-----N 159
QY	184 PSSALQQAALTCLKIRFENVHNDIFREIPGFQLETVKTLTLLPIYAQAANPHNLILQGAEL 243
DB	160 PHST--RSAALVKERFGNAELIRTNMGSFSQTNVETPLPTVQAASLHLLVWRDVQIY 217
QY	244 ADEWNADIHPSQIBENAGTSDDYKLLKENIPKYSNYCANTYRTGLKNLRDEPNMKWSIF 303
DB	218 GKEWG---YPO-----NDIDLFYKEQVSYTARYSDHCQWYNAGLNKLGRTGAKQWVDY 268
QY	304 NDVRYWTITVLDTISQSLVDYDKRYDSIGGIEVKGIKNELTREIYVTTREIN----- 355
DB	269 NRFREMNWDLVALFPNDARY-----PLETNA---ELTREIFTDPGVSVITGQSS 320
QY	356 -----FDRLPOLRQVPQNIATMEYNLTRASFKLFSLEQFIYFT-----ENTNF--GN 400
DB	321 TLISWYDMIP--AALPSPSTLENLLRKPDF--FTLLQEIWMYTSFRQNGTIEYYNMGWGQ 376
QY	401 RLVCISNRDAPTYNTITETLYGERTGSPTTKIRPFESYKYSIVTDRQSPVPYSPIQPHF 460
DB	377 RLTI-----LSYITYGS-----SFNKYSYGLAGAEIDIPGVQNDIYR 411
QY	461 I-----INOIELYANGSSNNTLK-YS-----AGG--SLSNYQNTITFFQFPR 498

A;Cross-references: UNIPARC:UPI000002B8CD
A;Experimental source: var. Tenebrionis
R;Adams, L.F.; Mathewes, S.; O'Hara, P.; Petersen, A.; Guertler, H.
Mol. Microbiol. 14, 381-389, 1994
A;Title: Elucidation of the mechanism of CryIIIA overproduction in a mutagenized strain
A;Reference number: S60781; MUID:95131759; PMID:7830581
A;Accession: S60781
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 9-652 <ADA>
A;Cross-references: UNIPARC:UPI000002B8CD; EMBL:U10985; NID:9506182; PIDN:AA432866.1; PI
A;Experimental source: var. tenebrionis strain NB176, a mutant of strain NB125
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1994
R;Wu, S.J.; Dean, D.H.
J. Mol. Biol. 255, 628-640, 1996
A;Title: Functional significance of loops in the receptor binding domain of Bacillus th
A;Reference number: S62317; MUID:96163559; PMID:8568902
A;Contents: annotation
R;Donovan, W.P.; Gonzalez, J.M.
Mol. Gen. Genet. 214, 365-372, 1988
A;Title: Isolation and Characterization of EG2158, a new strain of Bacillus thuringiensis
A;Reference number: I39812; MUID:89112139; PMID:3146015
A;Accession: I39812
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 9-652 <RES>
A;Cross-references: UNIPARC:UPI000002B8CD; GB:M37207; NID:g142735; PIDN:AAA50255.1; PID:
A;Experimental source: strain EG2158
R;Teixeira De Souza, M.; Lecadet, M.M.; Lereclus, D.
J. Bacteriol. 175, 2952-2960, 1993
A;Title: Full expression of the cryIIIA toxin gene of Bacillus thuringiensis requires a
A;Reference number: I39813; MUID:93259939; PMID:8491716
A;Accession: I39813
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 9-58 <RE2>
A;Cross-references: UNIPARC:UPI00000B007B; GB:L03393; NID:g304150; PIDN:AAA22350.1; PID:
A;Gene: cryIIIA
C;Superfamily: Parasporal crystal protein
C;Keywords: delta-endotoxin

Query Match 15.3%; Score 592.5; DB 2; Length 652;
Best Local Similarity 26.4%; Pred. No. 2.9e-31;
Matches 196; Conservative 101; Mismatches 307; Indels 139; Gaps 27;

QY 1 MNQNDNNEVEIIDSHTSPYFPRNNSDRYPYTNPNQPLQNTYKWLNMCOGNTQYG 60
DB 9 MPPNN-RSEHDTIKITENNEVP---TNHVQYPLAETPNPTLDNLYKEFLRMATDN--- 60

QY 61 DNFETPASADTTAAVSAGTIVSGTLLAGIGGLTSISGPIIGIIGIISFTGLITVFWPAG 120
DB 61 -NTEALDSSTTKDVIQKISVVGDLGVVG-----FPFG--GALVSFYTNFLTITWPE 111

QY 121 EODKTWTFQIKMGEIFVDTPTTESIKQLKLOTLEGFRQILQSYNTALDDWRKLRLOAP 180
DB 112 DP-----WKAFMEQVEALMDQKIADYAKNRKALAEQLQNNVEDVYSAWSQK----- 160

QY 181 GLPPSALQQAALTALKIRENVHNDIREIPGQLETYKTLILLPIYAAQANFHLNLLQOG 240
DB 161 -NPVSSRNPHSQRIELFSQAESHFRNSMPSPAISGYEVLFLTYYAQAANTHLFLKDA 219

QY 241 AELADEW---NADIHSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKNLRDEPN 297
DB 220 QYGEWGEWKEKEDI-----AEFYKRLKLTQBYTHCVKYNVGLDKLRGSSY 267

QY 298 MKWISFNDYRRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKIKNELREIYVTEI--- 354
DB 268 ESWVFNRRVREMTLVLDLIAFLPYLDVRLY-----PKEVKTELTRDVLTDPIGV 319

QY 355 -----NFDRLPQLRQVQPNLATMEYNLATRASFKLSFLEQFIYTE----- 394
DB 320 NNLRGYGTTFNSNIENIRKPH-----LPDYLHRIQFTRFPQGY 359

RESULT 11

S00873

Parasporal crystal protein cryBa1 - Bacillus thuringiensis subsp. thuringiensis

N;Alternate names: parasporal crystal protein cryA4

C;Species: Bacillus thuringiensis subsp. thuringiensis

C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 05-Oct-2004

C;Accession: S00873

R;Brizzard, B.L.; Whiteley, H.R.

Nucleic Acids Res. 16, 2723-2724, 1988

A;Title: Nucleotide sequence of an additional crystal protein gene cloned from Bacillus t

A;Reference number: S00873; MUID:88203216; PMID:3362680

A;Accession: S00873

A;Molecule type: DNA

A;Residues: 1-1228 <BRI>

A;Cross-references: UNIPROT:P05517; UNIPARC:UPI0000126BDC; EMBL:X06711; NID:g40264; PIDN:

C;Genetics:

A;Gene: cryA4

A;Start codon: TTG

C;Superfamily: Parasporal crystal protein

C;Keywords: delta-endotoxin

Query Match 14.9%; Score 576; DB 2; Length 1228;

Best Local Similarity 25.6%; Pred. No. 9.4e-30;

Matches 202; Conservative 107; Mismatches 285; Indels 196; Gaps 32;

QY 2 NONNDNNEVEIIDSHTS--PYFPRNNSDRYPYTNPNQPLQNTYKWLNMCOGNTQY 59

DB 4 NRKNENIINAVSNHSAQMDLLPDARIEDS-----LCIAEGN--- 40

QY 60 GDMFETPASADTTAAVSAGTIVSGTLLAGIGGLTSISGPIIGIIGIISFTGLITVFWPA 119

DB 41 --NIDFPVSAST---VQTGINIAGRILGVLG-----VPPAQLASFYSLVGLWPR 87

QY 120 GEODKTWTFQIKMGEIFVDTPTTESIKQLKLOTLEGFRQILQSYNTALDDWRKLRLOA 179

DB 88 GRDQ---WEIFLHEVEQLINQITENARNALTARLQGLGDSFRAYQQSLEDWLE---NR 140

QY 180 PGLPPSALQQAALTALKIRENVHNDIREIPGQLETYKTLILLPIYAAQANFHLNLLQO 239

DB 141 DDARTSRLVLTQYIALEL-----DFINAMPLFAIRNQEVPLLMVYAAQANLHLLLRD 193

QY 240 GAELADEWNADIHPSQIEPNAGTSD--YKLLKENIPKYSNYCANTYRTGLKNLRDEP 296

DB 194 ASLFGSEFGL-----TSQEIQRYEROVERTRDYSDYCVIEWYNTGLSLRGTN 241

QY 297 NMKWSIFNDYRRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKIKNELREIYVTEIINF 356

242 AASWRVYNQFRDRLTLGVLDLVALPSPVDRTYPTINTSA-----QLTRVYTDIGA 293

357 DRLPQLRVQNLATWE-YNLTRASFK-----LFSELEQFIYFTENTNFGNRLVG 404

294 TGV-----NMASMNWNNAPSFSALEAAAIKSPHLLDLEQLTIFSASSRW----- 340

405 ISNRDARTY--SWTITETLYG-----ERTGSPPTTKTIRPF-----ESYKVSIVT 446

341 -SNTRHMTYWRGHTIQSRPIGGGLNTSTHGATNTSINPVTILRFASRDVYRTESYAGVLLW 399

447 DRQSPPVVS--PIOPHFIIQIELYLNGSSNNLTKXSAGGSLSNQNTTFFQPPRKDCNL 504

400 GYILEPIHGVTVRFTNPNQISDRGTANYSQYSPGL-----QLKQSET 446

505 VIDPGCS--PNFNMYSHILSHFLSFTYSYVIGLQQL-LDTGVLGWTHSSVDRNAISDK 561

447 ELPEPTTERPNYESYRSLH-----IGIILQSRVNVVYSWTHRSADRTNTIGPN 497

562 IITMIPAIGNLDTNSKVIIEGPGHGTGNLVYLSQ---GRLEITCTPTNSTQSYFIRLR 618

498 RITQIPMVKASELPQGTTVVRGPGFTGGDILRRNTGCGFPIRVTVNGP-LTQYRIGFR 556

619 YAT-----NGAGNTLPLNLSITIPGVIIPPORLANTFSGTYNNLQVGDGFGYFPPS 670

557 YASTVDPDFVSRGGTIVNNFR-----LFTMNSG---DELKYGNFVYRAFT 601

671 TVTLPLNRNIPFPFNR-ADV-----SNSILIIDKIEFIPITSMHQWREKQKLETI 720

602 -----PFTFTQIQDIIRTSIQLSGNGEVYIDKILIPVTAFEAYD---LERA 648

721 QTKINTPFTN 730

649 QEAVALFTN 658

RESULT 12

I39815

insecticidal protein cryv - *Bacillus thuringiensis*

C;Species: *Bacillus thuringiensis*

C;Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 05-Oct-2004

C;Accession: I39815

R;Gleaves, A.P.; Williams, R.; Hedges, R.J.

Appl. Environ. Microbiol. 59, 1683-1687, 1993

A;Title: Screening by polymerase chain reaction of *Bacillus thuringiensis* serotypes for iensis subsp. kurstaki.

A;Reference number: I39815; MUID:93298009; PMID:8517758

A;Accession: I39815

A;Status: preliminary; translated from GB/EMBL/DBD

A;Molecule type: DNA

A;Residues: 1-719 <RES>

A;Cross-references: UNIPROT:Q45752; UNIPARC:UPI0000036008; GB:M98544; NID:gl42767; IDN

C;Genetics:

A;Gene: cryv

C;Superfamily: Parasporal crystal protein

Query Match 14.7%; Score 568.5; DB 2; Length 719;

Best Local Similarity 26.2%; Pred. No. 1.3e-29;

Matches 208; Conservative 107; Mismatches 283; Indels 195; Gaps 35;

QY 3 QNNDNNEYEIIDGHTSPYFPRNNSDRYPYTNPNQPLQNTNYKEWLNMCQGNQTYQDN 62

DB 4 KNQDKHQ-----SFSSNAKVDKISTDS---LKNETDIELQNHEDCLKM---SEY-EN 50

QY 63 FEFASADTTAAVSAGTIVSGTLLAGIGLTSISGPIGI--IGNAILISFGTLITVFWPAG 120

DB 51 VEFVVSASTIQ-----TGIIAGIKILGTIGVPFAGQVASLYSFIILGELPKG 97

QY 121 EQDKTVMTQPIKMGIEIFVDTPLTYESIKQLKIQLEGFQRIQLQSYNTALDDW---RKLKRL 177

DB 98 ---KNQWEIPEWHEVEEIIQKISTYARKALTDLKGIDALAVYHDSLEWVGNNNTRA 154

QY 178 QAPGLPSSALQQAALTKIRFENVHNDFFIREPGFQLETYKTLTLLPIYAAANFHLNLL 237

Db	155	R-----SVVKSQVIALELM-----FVQKLPSFAVSGEEVPLLPYIAQAANLHLLLL	200
Qy	238	QOGAELADEWNADIHPSOLEPNAGTSDDYVKLLKENIPKYSNYCANTYRTGLKNLRDEPN	297
Db	201	RDASIFGKEWG--LSSEI-----STFYNQVERAGDYSDHCVKWYSTGLNLRGTNA	251
Qy	298	MKWSIFNDYRYMTITVLDTISQFSQLDIDIKRYRDSIGGIEVKGIKN--ELTREIYITTEIN	355
Db	252	ESWRYNQFRDQMTLWLDLVALFSYDTQMY-----PIKTTAQLTREYVT-----	297
Qy	356	FDRLPOLRVQPNL-ATMEYNLTRAFK-----LFSFLEQFIFYT-----ENTNF	398
Db	298	-DAIGTVHPHPSFTTWTNNNAPSFAIEAAVVRNPHLLDLEQVTIYSLLSRWSNTQY	356
Qy	399	GNRLVGISNRDAPTSNITETLYGERTGSPITTKIRPEES---YKVSIVTDQK---SPP	452
Db	357	MN-MWGGHKLFRFTGGTGLNISTQGSTNTSINPVTL-PFTSRDVRKTESLAGLNLFQTQ	414
Qy	453	VSPIQP-----HFIINQI---ELYLNGSSNNLTKYSGAGSLSNYQNTTFFQPRKKDCNL	504
Db	415	VNGVPRVDHMKFVTHPIASDNFYPGYA-----GIGTQLQDSN-----EL	456
Qy	505	VIDPCCSNFNFNYSHILSHFLSYFTVYVIGLOQLDTCVLGWTHTSSVDRYNAISKIIT	564
Db	457	PPEATGQPNYESYSHRLSHIGLISASHVKAL-----VYSWTHRSADRTNTEPNSIT	508
Qy	565	MIPAIKGNLDTNSKVIKPGHGTGCVLAVYLSQG-----RLEITCETPNSTQSYFIRLRY	619
Db	509	QIPLVKAFLNSGAARVGPQGTGGDILRRNTGTGDIRVNI---NPPFPAQRYRVIRY	565
Qy	620	AT-----NGAGNTLPNLSITPIGVIGIPPORLANTFSGTNNYNNIQDGFYGFQFP	669
Db	566	ASTTDLQPHTSINGKAINOGNFSATM-----NRGEDLDYKTFRTVGFT	608
Qy	670	STVTLPNLRNPFFPNRADVNSILI-----IDKIEFIPITSSMHQNRKQKL	717
Db	609	T-----PFSP--LDVQSFTTIGAWNFSSGNEGVNIDRIEFVPEVTEAYD---F	653
Qy	718	ETIQTKINTFTFN	730
Db	654	EKAQEKVTALFTS	666
RESULT 13			
I39814			
insecticidal protein cryv1 - Bacillus thuringiensis			
C:Species: Bacillus thuringiensis			
C>Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 05-Oct-2004			
C:Accession: I39814			
R:Shin, B.S.; Park, S.H.; Choi, S.K.; Koo, B.T.; Lee, S.T.; Kim, J.I.			
Appl. Environ. Microbiol. 61, 2402-2407, 1995			
A:Title: Distribution of cryV-type insecticidal protein genes in Bacillus thuringiensis & tomoscidus.			
A:Reference number: I39814; PMID:95314293; PMID:7793960			
A:Accession: I39814			
A:Status: preliminary; translated from GB/EMBL/DBJ			
A:Molecule type: DNA			
A:Residues: 1-719 <RES>			
A:Cross-references: UNIPARC:UPI0000036009; GB:J36338; NID:G540281; PIDN:AAC36999.1; PID:G			
C:Genetics:			
A:Gene: cryv1			
C:Superfamily: Parasporal crystal protein			
Query March 14.6%; Score 565.5; DB 2; Length 719;			
Best Local Similarity 26.2%; Pred. No. 2.1e-29;			
Matches 208; Conservative 106; Mismatches 284; Indels 195; Gaps 35;			
Qy	3	QNNDNNEIIDSHTSPYFPNRRNSDRPYTPNPNQPLQNTNYKEWLNMCQNTQYGDN	62
Db	4	KNQKHQ-----SPSSNAKVDKISTDS---LKNETDIELQNINHEDCLW-----SEY-EN	50
Qy	63	FETFASADTTAAVSGAGTIVSGTLLAGLGLTSISGPIGI--ICAILISFGTLITVFWPAG	120

```
Db 51 VEPFVSASTIQ-----TGIGIAGKILGTGLGVPFAGQVASLSYFILGELWPKG 97
QY 121 EODKTVTQFKMGIEIFVDTPLTESIQLKLOTLEGFQRILOSNTALDDW---RKLKRL 177
Db 98 ---KNQWEIEMEHVEEIIINQKISTYARNKALTDLKGLDALAVYHDSLESVWGNRNTRA 154
QY 178 QAPGLPSSALQQAALTKIRFENVHNDIFREIPGFOLETYKTLPIYAQAANFHLNLL 237
Db 155 R-----SVVKSOYIALELM-----FVKLSPFAVSAGEVPLLPYIAQAANLHLLL 200
QY 238 QQAELADEWNADIHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKNLRDEPN 297
Db 201 RDAISIFGKEWG--LSSEI-----STFYNQVERAGDYSYHCVKWYSTGLNLRGTNA 251
QY 298 MKSIFNDYRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKGIKN--ELTREIYTTTEIN 355
Db 252 ESWRYNQFRDRTLMLVDLVALFPSYDTQMY-----PIKTTAQLTREYVT----- 297
QY 356 FDRLPQLRVQPNL-ATMEYNLTRASFK-----LPSLFQFIFYT-----ENTNF 398
Db 298 -DAIGTVHPSPFTSTTWYNNAPSFSAIEAAVVRNPHLLDFLEQVTIYSLLSRWSTQY 356
QY 399 GNLVIGISNRDAPTSNTITETLYGERTGSPPTKTIRPFES---YKVSIVTDRO---SPP 452
Db 357 MN-MWGGHKLFEFTIGTGLNISTQGSTNTSINPVTL-PFTSRDVRVYTESLAGLNFLTP 414
QY 453 VSPIQP-----HFIINQI---ELYNGSSNNTLKYSAGGSLSNYONTTFFQPRKKDCNL 504
Db 415 VNGVPRVDFHFKFVTHPIASDNFYYPGYA-----GIGTQLQDSEN-----EL 456
QY 505 VIDPGCSPFNNSYSHLSHFSLFYTVYVIGIQLQILDGTGLVGTWTHSSVDRYNAISDKIIT 564
Db 457 PPEATGQPNYESYSHRLSHIGLISASHVKAL-----VYSWTHRSADRTNTEPNSIT 508
QY 620 AT-----NGAGNTLPNISLTIPGVIGIPQRLNNTFSGTNNYNNLOYGDFGYFQPP 669
Db 566 ASTTDLQFHTSINGKAINQGNFSATM-----NRGEDLDYKTFRTVGFT 608
QY 670 STVTPLPLNRNIPPIFNRAVDVNSILI-----IDKIEFIPITSSMHQNRKQKL 717
Db 609 T-----PFSP--LDVQSTFTIGAWNFSSGNEVYIDRIEFPVEVTYAEYD---F 653
QY 718 ETIQTINTFTFN 730
Db 654 EKAQEKVTALFTS 666

RESULT 14
S25383
parasporal crystal protein cryIIal - Bacillus thuringiensis
N;Alternate names: delta-endotoxin; parasporal crystal protein cryV
C;Species: Bacillus thuringiensis
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 05-Oct-2004
C;Accession: S25383
R;Tailor, R.; Tippet, J.; Gibb, G.; Pells, S.; Pike, D.; Jordan, L.; Ely, S.
Mol. Microbiol. 6, 1211-1217, 1992
A;Title: Identification and characterization of a novel Bacillus thuringiensis delta i-end
A;Reference number: S25383; MUID:92269582; PMID:1588820
A;Accession: S25383
A;Molecule type: DNA
A;Residues: 1-719 <TAI>
A;Cross-references: UNIPROT:Q45752; UNIPARC:UPI0000002DB73; EMBL:X62821; NID:g40289; PIDN
C;Genetics:
C;Superfamily: Parasporal crystal protein
C;Keywords: delta-endotoxin

Query Match 14.6%; Score 565.5; DB 2; Length 719;
Best Local Similarity 26.2%; Pred. No. 2.1e-29;
```

```
Matches 208; Conservative 106; Mismatches 284; Indels 195; Gaps 35;
QY 3 QNNDNNEVEIIDSHTSPPFPNRSNDSRYPTNNPNQPLONTNYKWLNNCQGNQYQGDN 62
Db 4 KNQDKHQ-----SFSSNAKVDKISTDS---LKNETDIELQNIHEDCLKN---SEY-EN 50
QY 63 PETFASADTIAAVSAGTIVSGTLLAGIGLGTISGPIGI--IGAIISFGTFLITVFWPAG 120
Db 51 VEPFVSASTIQ-----TGIGIAGKILGTGLGVPFAGQVASLSYFILGELWPKG 97
QY 121 EODKTVTQFKMGIEIFVDTPLTESIQLKLOTLEGFQRILOSNTALDDW---RKLKRL 177
Db 98 ---KNQWEIEMEHVEEIIINQKISTYARNKALTDLKGLDALAVYHDSLESVWGNRNTRA 154
QY 178 QAPGLPSSALQQAALTKIRFENVHNDIFREIPGFOLETYKTLPIYAQAANFHLNLL 237
Db 155 R-----SVVKSOYIALELM-----FVKLSPFAVSAGEVPLLPYIAQAANLHLLL 200
QY 238 QQAELADEWNADIHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKNLRDEPN 297
Db 201 RDAISIFGKEWG--LSSEI-----STFYNQVERAGDYSYHCVKWYSTGLNLRGTNA 251
QY 298 MKSIFNDYRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKGIKN--ELTREIYTTTEIN 355
Db 252 ESWRYNQFRDRTLMLVDLVALFPSYDTQMY-----PIKTTAQLTREYVT----- 297
QY 356 FDRLPQLRVQPNL-ATMEYNLTRASFK-----LPSLFQFIFYT-----ENTNF 398
Db 298 -DAIGTVHPSPFTSTTWYNNAPSFSAIEAAVVRNPHLLDFLEQVTIYSLLSRWSTQY 356
QY 399 GNLVIGISNRDAPTSNTITETLYGERTGSPPTKTIRPFES---YKVSIVTDRO---SPP 452
Db 357 MN-MWGGHKLFEFTIGTGLNISTQGSTNTSINPVTL-PFTSRDVRVYTESLAGLNFLTP 414
QY 453 VSPIQP-----HFIINQI---ELYNGSSNNTLKYSAGGSLSNYONTTFFQPRKKDCNL 504
Db 415 VNGVPRVDFHFKFVTHPIASDNFYYPGYA-----GIGTQLQDSEN-----EL 456
QY 505 VIDPGCSPFNNSYSHLSHFSLFYTVYVIGIQLQILDGTGLVGTWTHSSVDRYNAISDKIIT 564
Db 457 PPEATGQPNYESYSHRLSHIGLISASHVKAL-----VYSWTHRSADRTNTEPNSIT 508
QY 565 MIPAIGNNLDTNSKVIEGPGHTGGLNLYVLSQSG-----RLIECTEPNSTQSYFIRLY 619
Db 509 QIPLVKAFLSSGAAVVRGPGFTGDIILRTNTGTGDIRVNI---NPPFAQRYRVI 565
QY 620 AT-----NGAGNTLPNISLTIPGVIGIPQRLNNTFSGTNNYNNLOYGDFGYFQPP 669
Db 566 ASTTDLQFHTSINGKAINQGNFSATM-----NRGEDLDYKTFRTVGFT 608
QY 670 STVTPLPLNRNIPPIFNRAVDVNSILI-----IDKIEFIPITSSMHQNRKQKL 717
Db 609 T-----PFSP--LDVQSTFTIGAWNFSSGNEVYIDRIEFPVEVTYAEYD---F 653
QY 718 ETIQTINTFTFN 730
Db 654 EKAQEKVTALFTS 666

RESULT 15
JH0261
parasporal crystal protein cry3Ca1 - Bacillus thuringiensis subsp. kurstaki (strain BT10)
N;Alternate names: parasporal crystal protein cryIIID
C;Species: Bacillus thuringiensis subsp. kurstaki
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 05-Oct-2004
C;Accession: JH0261; S18944
R;Lambert, B.; Theunis, W.; Aguda, R.; Van Audenhove, K.; Decock, C.; Janssens, S.; Seuri
Gene 110, 131-132, 1992
A;Title: Nucleotide sequence of gene cryIIID encoding a novel coleopteran-active crystal
A;Reference number: JH0261; MUID:92184108; PMID:1544571
A;Accession: JH0261
A;Molecule type: DNA
A;Residues: 1-649 <LAM>
```

Search completed: December 4, 2005, 13:12:34
Job time : 45 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2005, 12:51:52 ; Search time 230 Seconds
(without alignments)
2254.623 Million cell updates/sec

Title: US-10-783-417-2
Perfect score: 3869
Sequence: 1 MQQNDNVEIIDSHTSPY.....KLETIQTKINTFFNHTKTL 735

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1060.5	27.4	1180	1	P16480 bacillus th
2	1060.5	27.4	1180	2	Q7AL67 BACTI
3	892.5	23.1	675	1	C10AA BACTI
4	892.5	23.1	675	2	Q8KNV2 BACTI
5	862.5	22.3	1128	2	Q9FDC0 BACTF
6	849	21.9	1109	1	C28AA BACTF
7	847	21.9	650	2	Q8VNX2 BACTV
8	803.5	20.8	683	2	Q75VA2 BACTE
9	784	20.3	688	2	Q8VNX1 BACTV
10	775.5	20.0	1169	1	CR8BA BACUK
11	758.5	19.6	1136	1	CR4BA BACTI
12	758.5	19.6	1136	2	Q7AL72 BACTI
13	734	19.0	688	2	Q5W7N9 BACTU
14	709	18.3	682	1	C19BA BACUH
15	705	18.2	829	2	Q6BE06 BACTU
16	677	17.5	825	2	Q6BE09 BACTU
17	676	17.5	659	1	CR3BA BACTO
18	672.5	17.4	648	1	C19AA BACTU
19	663	17.1	1169	2	Q56B08 BACTU
20	662	17.1	660	2	CR0U6 BACTA
21	660.5	17.1	1138	1	CR7AB BACUK
22	654	16.9	652	1	CR3BB BACTU
23	649	16.8	686	2	Q75Q05 BACTE
24	641.5	16.6	1138	1	CR7AA BACTU
25	639.5	16.5	1280	2	Q8VUK9 BACTU
26	633	16.4	1144	2	Q8KZ17 BACTG
27	632.5	16.3	1138	1	CR7AB BACUA
28	630	16.3	1236	2	Q939T3 BACTU
29	619	16.0	1231	1	Q8KNY2 BACTU
30	618.5	16.0	1169	1	CR9DA BACTP
31	618	16.0	1160	1	CR8CA BACTP

32	618	16.0	1160	2	Q6R2R6 BACTU	Q6r2r6 bacillus th
33	615.5	15.9	1157	1	CR8AA BACUK	Q45704 bacillus th
34	615	15.9	1231	1	CR1BD BACTZ	Q9zaz5 bacillus th
35	612.5	15.8	675	2	Q6BCH5 BACTU	Q6bch5 bacillus th
36	612	15.8	1157	1	CR9CA BACTO	Q45733 bacillus th
37	604	15.6	1163	2	Q5XLA8 BACTP	Q5xla8 bacillus th
38	599	15.5	826	1	C27AA BACUH	Q9s597 bacillus th
39	597	15.4	1340	2	Q589X2 PARPP	Q589x2 paenibacill
40	595	15.4	1344	2	Q76SX7_9BACL	Q76sx7 paenibacill
41	594.5	15.4	652	2	Q6PXN8 BACTU	Q6pxn8 bacillus th
42	592.5	15.3	644	1	CR3AA BACTM	P0a381 bacillus th
43	592.5	15.3	644	1	CR3AA BACTM	P0a380 bacillus th
44	592.5	15.3	644	1	CR3AA BACTT	P0a379 bacillus th
45	592.5	15.3	652	2	Q9S6N9 BACTU	Q9s6n9 bacillus th

ALIGNMENTS

RESULT 1
CR4AA_BACTI STANDARD; PRT; 1180 AA.
AC P16480;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Pesticidal crystal protein cry4Aa (insecticidal delta-endotoxin
DE CryIVA(a)) (Crystalline entomocidal protoxin) (135 kDa crystal
DE protein).
GN Name=cry4Aa; Synonyms=cryIVA(a), isrH4;
OS Bacillus thuringiensis subsp. israelensis.
OG Plasmid 72 Kb.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=1430;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sen K., Honda G., Koyama N., Nishida M., Neki A., Sakai H., Himeno M.,
RA Komano T.;
RT "Cloning and nucleotide sequences of the two 130 kDa insecticidal
RT protein genes of Bacillus thuringiensis var. israelensis.";
RL Agric. Biol. Chem. 52:873-878 (1988).
[2]
RN NUCLEOTIDE SEQUENCE.
RA MEDLINE=88015571; PubMed=2821500;
RX Ward E.S., Ellar D.J.;
RT "Nucleotide sequence of a Bacillus thuringiensis var. israelensis gene
RT encoding a 130 kDa delta-endotoxin.";
RL Nucleic Acids Res. 15:7195-7195(1987).
[3]
MUTAGENESIS STUDIES.
RX MEDLINE=94307434; PubMed=7913448; DOI=10.1016/0014-5793(94)00604-0;
RA Nishimoto T., Yoshiaue H., Ihara K., Sakai H., Komano T.;
RT "Functional analysis of block 5, one of the highly conserved amino
RT acid sequences in the 130-kDa CryIVA protein produced by Bacillus
RT thuringiensis subsp. israelensis.";
FEBS Lett. 348:249-254(1994).
CC -!- FUNCTION: Promotes colloidotomic lysis by binding to the midgut
CC epithelial cells of insects.
CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -!- MISCELLANEOUS: Diverse amino acid mutations in sequence block 667-
CC 676 have no direct effect on the insecticidal activity but alter
CC the structural stability of the toxin protein molecule.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not


```
CC removed.
CC EMBL; D00248; BAA00179.1; -; Genomic DNA.
DR EMBL; Y00423; CAA68448.1; -; Genomic DNA.
DR PIR; A26858; A26858.
DR PIR; I39870; I39870.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR Pfam; PF03944; Endotoxin_C; 1.
DR Pfam; PF00555; Endotoxin_M; 1.
DR Pfam; PF03945; Endotoxin_N; 1.
KW Plasmid; Sporulation; Toxin.
FT CONFLICT 306 306 L -> V (in Ref. 2).
FT CONFLICT 1052 1052 T -> I (in Ref. 2).
FT CONFLICT 1109 1109 R -> G (in Ref. 2).
FT CONFLICT 1127 1127 C -> W (in Ref. 2).
SQ SEQUENCE 1180 AA; 134539 MW; 6FB5B6979DACAD3B CRC64;

Query Match 27.4%; Score 1060.5; DB 1; Length 1180;
Best Local Similarity 35.1%; Pred. No. 1.1e-59;
Matches 271; Conservative 118; Mismatches 278; Indels 105; Gaps 24;

QY 1 MNQNDNNEVEIIDSHTSYFPNRRNSDRPYTNPNQPLQNTNYKEWLNMCQNTQYG 60
DB 1 MNPYQNKNEYETLNASQKKL--NISNNYTRYPIENSPKQLQSTNYKDWLNMCQNTQYG 58

QY 61 DNFETPASADTIAAASAGTIVSGTLGAGIGLTSISGPIGIIAGIISFGTLTVFPWAG 120
DB 59 GDFETFIDS--GELSAIYIVGTVLTGFGFT---PLGL---ALIGFGLTLPVLPFAQ 108

QY 121 EQDKTWTQFIKMGEIFVDTPLETESIKQLKQLEGRQILOSYNTALDWRKLRLOAP 180
DB 109 DQSNLT--WSDFITQTKNIKKEIASTVISNANKILNRSFNVISTYHNHLKTWE---NNP 162

QY 181 GLPPSSALQQAALTALKIRFENVHDFIREIP--GFQLETYKTLILLPIYAOAANFHLNLQ 238
DB 163 NPQNTQDVRTQIQLVHYHFQNVIPELVNSCPPNPSCDYNNILVLSYAAQANLHLTVLN 222

QY 239 QGAELADEWNADIHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKNLRDEP-- 296
DB 223 QAVKEAYLKNRRQFDYLEP-LPTAIDYYPVLTKAIEDYNYCVTVTKYKGLNLKTTTDS 281

QY 297 ----NMKWSIFNDYRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKIGKNELTREIYTT 352
DB 282 NLDGNINWNTYNTYTKMTTAVLDVALFPNYDVGY-----PIGVQSELTRIIYQV 333

QY 353 EINFDRLPQLRVQPNLATMEYNLTRASFLEQFIY-----TENTNFGNR----- 401
DB 334 -LNFESPYYKY--DFQYQEDSLTRRP-HLFTWLDLSLNFYEKAQTTPNNFTSHYNNFHY 389

QY 402 -LVGISNRDAPYTSNTITETLYGERGSPPTKTRIRPFESYKYSIVTDROSPPVSPQHPF 460
DB 390 TLDNISQKSSVFGNHNVTDKL--KSLGLATNIYI-----F 422

QY 461 IINQIEL---YLVGSSNNT-----LKYSAGGSLSNYQ-NTTFFQFPRKDC 502
DB 423 LLNVISLDNKNYLDNNYISKMDFFITNGTRLEKELTAGSQITYDVKNIIFGLPILKRR 482

QY 503 NLVIDPGCSNPNFNYSHILSHFSFTYSYVIGIQLQILDVGLGWTSHSSVDYRNAISDKI 562
DB 483 ENQGNPTLPPTYDNYSHILSFIKSLSPATYKTQVY-----TPAWTHSSVDPKNTIYTHL 537

QY 563 ITWIPAIKGNLDTNSKVIETGPGHTGGLNVLVLOSQGRLEITCETPNSTQSYFIRLRYATN 622
DB 538 TTIQIPAVKANSLGTAKSVQVGGHTGGDLI--DFKDHFKITCQHSFQOQSFIIRIYASN 595

QY 623 GAGNTLPNLSLTPGVIGIPQRLNNTFSGTNNLQYDGFYGFQFPSTVTLPLNRIPE 682
DB 596 GSANTRAVNLISIPGVAEL--GHALNPTFSGTDYTNLKYKDFQYLEFNEVKFAPNQNISL 654

QY 683 IFNRADV--SNSILIIDKIEFIPITSSMHQNRKQKLETIQTINTFTNHTK 733
DB 683 IFNRADV--SNSILIIDKIEFIPITSSMHQNRKQKLETIQTINTFTNHTK 733
```

```
DB 655 VFNRSDVYVTTVLIDKIEFLPITRSIREDEKOKLETVOQIINTFYANPIK 706

RESULT 2
Q7AL67 BACTI
ID Q7AL67 BACTI PRELIMINARY; PRT; 1180 AA.
AC Q7AL67
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Pesticidal crystal protein cry4AA.
GN Name-cry4AA; Synonyms-cryIVA(A), isrH4, p8t110;
OS Bacillus thuringiensis (subsp. israelensis).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=1430;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22235415; PubMed=12324359;
RX DOI=10.1128/AEM.68.10.5082-5095.2002;
RA Berry C., O'Neil S., Ben-Dov E., Jones A.F., Murphy L., Quail M.A.,
RA Harris D., Zaritsky A., Parkhill J.;
RT "Complete sequence and organisation of pBtoxis, the toxin-coding
RT plasmid of Bacillus thuringiensis subsp. israelensis.";
RL Appl. Environ. Microbiol. 68:5082-5095(2002).
DR EMBL; AL731825; CAD30148.1; -; Genomic DNA.
SQ SEQUENCE 1180 AA; 134538 MW; 6FB5B6979DACAD3B CRC64;

Query Match 27.4%; Score 1060.5; DB 2; Length 1180;
Best Local Similarity 35.1%; Pred. No. 1.1e-59;
Matches 271; Conservative 118; Mismatches 278; Indels 105; Gaps 24;

QY 1 MNQNDNNEVEIIDSHTSYFPNRRNSDRPYTNPNQPLQNTNYKEWLNMCQNTQYG 60
DB 1 MNPYQNKNEYETLNASQKKL--NISNNYTRYPIENSPKQLQSTNYKDWLNMCQNTQYG 58

QY 61 DNFETPASADTIAAASAGTIVSGTLGAGIGLTSISGPIGIIAGIISFGTLTVFPWAG 120
DB 59 GDFETFIDS--GELSAIYIVGTVLTGFGFT---PLGL---ALIGFGLTLPVLPFAQ 108

QY 121 EQDKTWTQFIKMGEIFVDTPLETESIKQLKQLEGRQILOSYNTALDWRKLRLOAP 180
DB 109 DQSNLT--WSDFITQTKNIKKEIASTVISNANKILNRSFNVISTYHNHLKTWE---NNP 162

QY 181 GLPPSSALQQAALTALKIRFENVHDFIREIP--GFQLETYKTLILLPIYAOAANFHLNLQ 238
DB 163 NPQNTQDVRTQIQLVHYHFQNVIPELVNSCPPNPSCDYNNILVLSYAAQANLHLTVLN 222

QY 239 QGAELADEWNADIHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKNLRDEP-- 296
DB 223 QAVKEAYLKNRRQFDYLEP-LPTAIDYYPVLTKAIEDYNYCVTVTKYKGLNLKTTTDS 281

QY 297 ----NMKWSIFNDYRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKIGKNELTREIYTT 352
DB 282 NLDGNINWNTYNTYTKMTTAVLDVALFPNYDVGY-----PIGVQSELTRIIYQV 333

QY 353 EINFDRLPQLRVQPNLATMEYNLTRASFLEQFIY-----TENTNFGNR----- 401
DB 334 -LNFESPYYKY--DFQYQEDSLTRRP-HLFTWLDLSLNFYEKAQTTPNNFTSHYNNFHY 389

QY 402 -LVGISNRDAPYTSNTITETLYGERGSPPTKTRIRPFESYKYSIVTDROSPPVSPQHPF 460
DB 390 TLDNISQKSSVFGNHNVTDKL--KSLGLATNIYI-----F 422

QY 461 IINQIEL---YLVGSSNNT-----LKYSAGGSLSNYQ-NTTFFQFPRKDC 502
DB 423 LLNVISLDNKNYLDNNYISKMDFFITNGTRLEKELTAGSQITYDVKNIIFGLPILKRR 482

QY 503 NLVIDPGCSNPNFNYSHILSHFSFTYSYVIGIQLQILDVGLGWTSHSSVDYRNAISDKI 562
DB 483 ENQGNPTLPPTYDNYSHILSFILKSLSPATYKTQVY-----TPAWTHSSVDPKNTIYTHL 537

QY 563 ITWIPAIKGNLDTNSKVIETGPGHTGGLNVLVLOSQGRLEITCETPNSTQSYFIRLRYATN 622
DB 538 TTIQIPAVKANSLGTAKSVQVGGHTGGDLI--DFKDHFKITCQHSFQOQSFIIRIYASN 595

QY 623 GAGNTLPNLSLTPGVIGIPQRLNNTFSGTNNLQYDGFYGFQFPSTVTLPLNRIPE 682
DB 596 GSANTRAVNLISIPGVAEL--GHALNPTFSGTDYTNLKYKDFQYLEFNEVKFAPNQNISL 654

QY 683 IFNRADV--SNSILIIDKIEFIPITSSMHQNRKQKLETIQTINTFTNHTK 733
DB 683 IFNRADV--SNSILIIDKIEFIPITSSMHQNRKQKLETIQTINTFTNHTK 733
```

Db 538 TTQTPAVKANS LGTAS KVVQGGHGTGGDLI--DFKDHFKITCOHNSFOQSYFIRIRYASN 595
 Qy 623 GAGNTLPNISLTIPGIVGIPQRPNNLTSGTNNYNLQYDFGYFOPSPSTVTLPLNRNIPF 682
 Db 596 GSANTRAVINLSIPGVAEL-GWALNPTSGTDYTNLKYDFQYLFESNEVKFAPNQNL 654
 Qy 683 IFNRADV-SNGLIIDIKIEFIPITSSMHQNRKQKLETIQTKINTFFTNHTK 733
 Db 655 VFNRSVDVYNTTVALIDKIEFLPITRSIRREKREKQKLETVQQIINTFYANPIK 706

RESULT 3
 C10AA_BACTI STANDARD; PRT; 675 AA.
 AC P09662;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Pesticidal crystal protein cry10Aa (insecticidal delta-endotoxin
 DE CryXA(a)) (Crystalline entomocidal protein) (78 kDa crystal protein).
 GN Name=cry10Aa; Synonyms=cryIVC, cryXA(a);
 OS Bacillus thuringiensis subsp. israelensis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
 OC Bacillus cereus group.
 NCBI_TaxID=1430;
 [1]
 RN NUCLEOTIDE SEQUENCE [GENOMIC DNA].
 RC STRAIN=ONR60A;
 RX MEDLINE=86223796; PubMed=3011746;
 RA Thorne L., Garduno F., Thompson T., Decker D., Zoumes M., Wild M.,
 RA Walfield A.M., Pollock T.J.;
 RT "Structural similarity between the lepidoptera- and diptera-specific
 RT insecticidal endotoxin genes of *Bacillus thuringiensis* subsp.
 RT 'kurstaki' and 'israelensis'";
 RL J. Bacteriol. 166:801-811(1986).
 CC -!- FUNCTION: Promotes colloidosmotic lysis by binding to the midgut
 CC epithelial cells of mosquitos. Active on *Aedes aegypti*.
 CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
 CC sporulation and is accumulated both as an inclusion and as part of
 CC the spore coat.
 CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
 CC terminus.
 CC -!- SIMILARITY: Belongs to the delta endotoxin family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL: M12662; AAA22614.1; -; Genomic_DNA.
 DR PIR; B29838.
 DR InterPro; IPR001178; Endotoxin.
 DR InterPro; IPR005638; endotoxin_C.
 DR InterPro; IPR005639; endotoxin_N.
 DR Pfam; PF031944; Endotoxin_C; 1.
 DR Pfam; PF00555; Endotoxin_M; 1.
 DR Pfam; PF03945; Endotoxin_N; 1.
 DR Sporulation; Toxin.
 KW SEQUENCE 675 AA; 77761 MW; 2A70011EEA7985F5 CRC64;
 SQ

Query Match 23.1%; Score 892.5; DB 1; Length 675;
 Best Local Similarity 31.9%; Pred. No. 4e-49;
 Matches 245; Conservative 126; Mismatches 267; Indels 129; Gaps 30

Qy 1 MNQNNNNEYEIDSHTSPYFPNNDSRYPTTNNPQLQNTNYKEWLNCCQNTQYG 60
 Db 1 MNPYQNKNEYEIFNAPSGP--SKSNNSRYPLANKPNQPLKNTNYKDLNVCQDNQQY 58

Qy 61 DNPETTFASADTIAVSGACTIVSGTLLAGIGGTSISGPIGIIIGALIIISFTLITVFWPAG 120
 Db 59 NNAGNFASSETIVGVSAGIIIVVGTMIG-----AFAAP--VLAAGIIISFTLLPFW-QG 109

```
DR GO: 00030435; P: sporulation; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR Pfam; PF03944; Endotoxin_C; 1.
DR Pfam; PF00555; Endotoxin_M; 1.
DR Pfam; PF03945; Endotoxin_N; 1.
SQ SEQUENCE 675 AA; 77759 MW; 4EB0E51AA0372FF1 CRC64;

Query Match      23.1%; Score 892.5; DB 2; Length 675;
Best Local Similarity 31.9%; Pred. No. 4e-49;
Matches 245; Conservative 126; Mismatches 267; Indels 129; Gaps 30;

QY 1 MNQNDNNYEIDSHTSYPFPRNNDGRYPYNNPNQPLQNTYKELNMCQGNQYQ 60
Db 1 MNPYQNKNEYEFNAPSNGF--SKSNYSRYPLANKPNQPLKNTYKDWLNVCQDNOQYQ 58
QY 61 DNFETASADTTAANVAGTIVSGTLAGLGLTSISGPIGIIICAIISFGTILTVFWPAG 120
Db 59 NNAGNFASSETIVGVSAGIIVVGTMLG-----AFAAP--VLAAGIISFGTILPFW-QG 109
QY 121 EQDKTWTQIKMGEIFVDTPLTE---SIKQLKLTLEGFRQLQSYNTALDDWRKLR 177
Db 110 SDPANVQDLNIG-----GRPIQEIKNLIINVTISVTPIKQLDKIQEFDKWEPA-- 163
QY 178 QAPGLPPSSALQOAAALTKIRFENVHN---DFIREIPGFOLEYTKTLPIYAAANFHL 234
Db 164 -----THANAKAVHDLFTLEPIIDKOLDMLKNNASYRIPT-----LPYAIQIATWHL 211
QY 235 NLLQGAELADWADNIHSPQIEPNAGTSDDYK-LLKENIPKSYNCANTYGTGLKNLR 293
Db 212 NLLKHAATYNIW---LQNGQINPSTFNSNYYQYGLKRIQBYTCIQTYNAGLTMR 268
QY 294 DEPNKMSIFNRYRMTITVLDTSIQPSLYDIKRYRDSIGGIEVGKIKNELTREIYTT 353
Db 269 TWTNATWNYTRYLEMTITVLDLIAIFPNYDPEK-----PIGVKSELIREVY-TN 319
QY 354 INFDRLPQLRVQPNLATMEYNLTRASFLEQIFSYFETNTN-----FQNRLL 402
Db 320 VNSDTF-----RTILENGLTR-NPTLFTWINGRFYTRNSRDILDVDYDFISFTGNQM 372
QY 403 VGISNRDAPTYNITETILYGERGT---SPTTKTIRPFESY-----KVSIVTDQRSPPVSP 455
Db 373 -----AFTNDRDRIIAGVHGNIIISQDTSKVPFVRNKEIDKVEIVRHREYSDI 424
QY 456 IQPHFIINGIYLGSSNNLKYSGGSL-SNYQNTTFQPRKKDCNLVIDPGCSNPF 514
Db 425 -----YEMIFFNSSEVFYSNNTIENNYKRTDSYMPKQTKN----- 464
QY 515 NNYSHLSHSLFTSYSVYGLQIOLDTGVLGWTHSSVDRYNAISDKIITMIPAIKGNL 574
Db 465 KEYGHTLSYIKTDNYIFSVVRERRV---AFSWHTSVDFQNTDLDNITQIHAKALKV 521
QY 575 DTNSKVIQPGHTGGNLVYQSGRLIEITCETPNSTQSYFIRLRYATNGAGNTLPNLSLT 634
Db 522 SSDSKIVKPGHTGGDLVILKDSMDPRVF-LKNVSQYQVRIYATNA-----PKTTVF 575
QY 635 IGVGIGIPQRLNNTSGTNYN--NIQYGDGFGYQFPSTV-----TLPLNRNIPFI 683
Db 576 LTGIDTISVE-LPSTTSRQNPATDITYADFGVYVTFPTVPNKTFEGEDFLMT-----L 629
QY 684 FNRADVSNILIDKIEFTPTSSMHONREKQLETIQTKINTFFTN 730
Db 630 YGTPNHSYNI-YIDKIEFIPITQSVLDYTEKQNIETKQIKVINDLFVN 675

RESULT 5
Q9FDC0 BACTF PRELIMINARY; PRT; 1128 AA.
AC Q9FDC0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
```

Db	189	WNADQHPSPMLKSGT---	YYDELLVYIEKYINYCTKTYHKLGNHLKSEKITDAYNTY	243
Qy	307	RRYMTITVLDTISOFSLYDIKRYDRSDIGGIEVKGIGIKNELTREIYVTEINFDRLPQLRVQP	366	
Db	246	RRNTLVLDLVAVFPFDIRF-----	PRGVELELTREYVTSLDHLTRPP-----	291
Qy	367	NLATWEYNLTRASPKLFSFLBQFIYTYENTWFGNRLVGISNRDAPYSNT---	ITETLYG 423	
Db	292	-----	GLEFTWLSDIELYTESVAEGDYLGI--RESKYVTCNQPFMTMKNIVG	335
Qy	424	ERTG-SPTTKTIRPESKYSIVTDROSPVPSPIQPHFIINQIELYLANGSSNNTLKYSAG	482	
Db	336	NTNRLSKQLITLLEGE-FWTHLSINRPQTITAGINKLYSLQIKIVFTTFKNDN--	EYQKN 392	
Qy	483	GSLSNY---QNTTFQFPFKKDCNCLVDPGGSPFNFNYSHLISHPESLFTYSYVIGLQLOI	539	
Db	393	FNUNNQEPQETT-----	NYPNDYGGG-NSQKFKNLISHPPLIIH-----	KLEF 435
Qy	540	LD-----	TGVLGWHSSHVDRYNAISDKIITMTIPAKIGNNLDTNKSVIEGPGHTGGMVLVYQ	595
Db	436	AEYFHSIFALGWHSHNSVNSQNLISSESVSTQIPLKAYEV-TNNSVIRGPGFTGGDLIELR	494	
Qy	596	SQGRLEITCETPNSTQSYFIRLRYATGAGNTLNPISLITIPQVIGIPQRLNNTFTSGTY	655	
Db	495	D--KCSIKCKA-SSKKYAIASLFAANNAIAVSDVGSAGVL-----	LQPTFSRKGK 545	
Qy	656	NN-----	LOYGDFGYFOFPSTVTLPLNRNIPFIINRAD--VSNSILIIKIEFIPITSSM	708
Db	546	NNFTIQDLNKYDFQYHTLLVDIELPESEIHHILKREDDYEEGVLLIDKLEFKPIDENY	605	
Qy	709	HONREKQKLEITQTKTFFTNHTKTL	735	
Db	606	---TNEMNLEKAKKAVNVLFINATNAL	629	
RESULT 7				
Qy	8VNX2	BACTV PRELIMINARY;	PRT;	650 AA.
Db	8VNX2	BACTV PRELIMINARY;	PRT;	650 AA.
DT	01-MAR-2002	(TrEMBLrel. 20, Created)		
DT	01-MAR-2002	(TrEMBLrel. 20, Last sequence update)		
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)		
DE	Cry29Aa	protein.		
GN	Name=cry29Aa;			
OS	Bacillus thuringiensis (subsp. medellin).			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;			
OC	Bacillus cereus group.			
OX	NCBI_TaxID=79672;			
ON	NCBI_TaxID=79672;			
RA	NUCLEOTIDE SEQUENCE.			
RA	Delecluse A., Orduz S.;			
RL	Submitted (DEC-1995) to the EMBL/GenBank/DBDJ databases.			
DR	EMBL; AJ751977; CAC80985.1; -;	Genomic DNA.		
DR	GO; GO:0016787; F:hydrolase activity; IEA.			
DR	GO; GO:0005102; P:receptor binding; IEA.			
DR	GO; GO:0006952; P:defense response; IEA.			
DR	GO; GO:0009405; P:pathogenesis; IEA.			
DR	GO; GO:0030435; P:sporulation; IEA.			
DR	InterPro; IPR001178; Endotoxin.			
DR	InterPro; IPR005638; endotoxin_C.			
DR	InterPro; IPR005639; endotoxin_N.			
DR	Pfam; PF03944; Endotoxin_C; 1.			
DR	Pfam; PF00555; Endotoxin_M; 1.			
DR	Pfam; PF03945; Endotoxin_N; 1.			
SQ	SEQUENCE	650 AA; 74435 MW; DOCB846377CC517B	CRC64;	
Query Match				
Best Local Similarity 21.9%; Score 847; DB 2; Length 650;				
Matches 230; Conservative 124; Mismatches 273; Indels 126; Gaps 22				
Qy	1	MNQNDNNEYEIIDSHTSPVPFPPNRNSDRPYPTNNPQPLONTNYKEWLKNC--	QGNTQ 58	
Db	1	MNFGQNKNEYILNAPGS--ISNIPNNYSKYPIANTLQNTQNTYKDWTCANIDNNLK	58	

Qy	59	YGDNFEFASADTTAAVSAGTIVSGTTLIAGIGCLTISISGPTIGII	GAILIISFOTLITVFWP	118
Db	59	SINPFEANLQNSLWGLFAITAAIASLLSAPITGGTSAAGTAIAAAI	-----IPILWP	111
Qy	119	AGEODKTWTFQIKNGELFVDTPLTESIKOLKQTLLEGPRQILQSVNTALDWRKLRQ	178	
Db	112	S--OBNLPLDKLLAISEATLYSFQDQVRVEDALTRLESKDSVKFENAF	TWIN-----	164
Qy	179	APGLPPSALQOAAULTKIRPENVHNDFIREIPGQLETYKTL	LLPIYAAQANFHLNLQ	238
Db	165	---NPNST--NTTVRERFQEVNGRPFVGSMAFPRAKNYBPILLSTV	AAARLHLHLR	217
Qy	239	QCAELADEWNADIIHPSQIEPNAGTSDYYKLLKENIPKYSVCANTYRTGLKNL	RDEPNM	298
Db	218	DGITYAEAKNLSRQDDM-----PDGLLYKGFNKYCNEYIEHCIKWYNES	LLK--SVGA	271
Qy	299	KWSIFNDYRRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKGIKNELREI	VTTEINPDR	358
Db	272	NWLEYNQYRFTLASVDVLSLFSYDPRLYKERL--SVEI	-----LTKLYDPI	324
Qy	359	LPQURVOPNLATMBYNLFRASFKLFSFLEQFIYFTENTFGNRLVGISNRD	APTYSNTII	418
Db	325	GISLEADESKYTLEPRTLFTQLYTLTFYSNIFVNYMGHTN-----	TYRYLSP	370
Qy	419	ETLYGERTGSTTKTIRPPESYKUSIVTDROSPPVSPIQPHIINOIE	LYLN-----	470
Db	371	DKIPAERSFGKQSSYID-----KVPVIPNDKS-----IYKIRAYD	NHNGLEFNVM	415
Qy	471	-----GSSNNTLKYSAGGSLNSYNTQTTFFQPRKKQCNLVDPGCS	PNFNYSHLHS--	523
Db	416	YFGWDEKDOIQLIIGSSSTEY-----IKNCT-----	HLRADVISHD	454
Qy	524	--FSLFTSYVYIGLQLILDTVGLGTHSSVDRYNAISDKLITMT	PAIKGNLNDTSKVI	581
Db	455	LDEKNKCYSF-----AWTSTTISLENEIKNDIIITQIPAVKAYQLG	VQSQVI	500
Qy	582	EGPHTGGLVYLOSQLEITCBLTPNSTQSYFIRLRYATNGANTUPLNIS	LTTPGVIG	640
Db	501	KGPHTGGDLLNLSKNSDYLRSCHOHLSNVTKKYFVRIRYATNGSLN	TRPIINITIP--	557
Qy	641	IPPOC--LNNTFSGTNYNLYQDGYQFQFSTVPLNRPINPIFENRADV	--SNSILIID	697
Db	558	MTPQGMVLNDFTSCTGYSNLEYQNGFYKEFLKEVTLPNFQISLSLT	NSILNLSQNSILIID	617
Qy	698	KIEFIPITSSMHQNRKQKLETIOTKINTFTFN	730	
Db	618	RIEFLPITPSIRKSKEOONLEKNOKTVNKLFPN	650	

RESULT 8	
Q75VA2_BACTE	
ID Q75VA2_BACTE PRELIMINARY;	PRT; 683 AA.
AC Q75VA2_	
DT 05-JUN-2004 (TrEMBLrel. 27, Created)	
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)	
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)	
DE Putative mosquito cidal toxin.	
GN Name=cry30Aa like;	
OS Bacillus thuringiensis (subsp. entomocidus).	
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;	
OC Bacillus cereus group.	
OX NCBI_TaxID=1436;	
RN [1]	
RP NUCLEOTIDE SEQUENCE.	
RC STRAIN=INA288;	
RA Ikuya T., Yamaya K., Ito T., Sahara K., Bando H., Asano S. ;	
RL Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.	
DR ENBL: AB125059; BAD00052.1; -; Genomic DNA.	
DR GO: G0:0016787; F:hydrolase activity; IEA.	
DR GO: G0:0005102; P:receptor binding; IEA.	
DR GO: G0:0006952; P:defense response; IEA.	
DR GO: G0:0009405; P:pathogenesis; IEA.	

[illegible]

RESULT 9	
Q8VNX1 BACTV	
ID Q8VNX1 BACTV PRELIMINARY;	PRT; 688 AA.
AC Q8VNX1;	
DT 01-MAR-2002	(TREMBLrel. 20, Created)
DT 01-MAR-2002	(TREMBlrel. 20, Last sequence update)
DT 01-MAR-2004	(TREMBlrel. 26, Last annotation update)

DE	Cry30Aa protein.	
GN	Names=Cry30Aa;	
OC	Bacillus thuringiensis (subsp. medellin).	
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;	
OC	Bacillus cereus group.	
OX	NCBI_TaxID=79672;	
RN	[1]	
RP	NUCLEOTIDE SEQUENCE.	
RA	Delecluse A., Orduz S.;	
RL	Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AJ251978; CAC80986.1; -; Genomic DNA.	
DR	GO; GO:0016787; F:hydrolase activity; IEA.	
DR	GO; GO:0005102; F:receptor binding; IEA.	
DR	GO; GO:0006952; P:defense response; IEA.	
DR	GO; GO:0009405; P:pathogenesis; IEA.	
DR	GO; GO:0030435; P:sporulation; IEA.	
DR	InterPro; IPR001178; Endotoxin.	
DR	InterPro; IPR005638; endotoxin_C.	
DR	InterPro; IPR005639; endotoxin_N.	
DR	Pfam; PF03944; Endotoxin_C; 1.	
DR	Pfam; PF00555; Endotoxin_M; 1.	
DR	Pfam; PF03945; Endotoxin_N; 1.	
SQ	SEQUENCE .688 AA; 77844 MW; 75057085BD7B3A9 CRC64;	
	Query Match . 20.3%; Score 784; DB 2; Length 688;	
	Best Local Similarity 32.1%; Pred. No. 4.4e-42;	
	Matches 237; Conservative 113; Mismatches 294; Indels 94; Gaps 30;	
QY	1 MNQNDNNEYIIDSHTSPYFPRNSDRPYTNPNQPLQNTYKRWLNMCQNTQYG 60	
DB	1 MNSYENKNEVEILD--TSQKNSMNSRYSKYPLTNPKVPLQNTYKDWLNMCQTITPLC 58	
QY	61 DNPETASADTTAAVAGTIVSGTLLAGIGLTSISGPIGIGIAIISFGTLITVFPAG 120	
DB	59 TPIDTDSKLVAIKVIGAI-----FKSMFGPGAAGVLKSFSTIIPILWP-- 105	
QY	121 EODKT-VWTOFIKWG-EIFVDTPLTESIK-----QLKLTLEGFRILQSYNTALDDWR 172	
DB	106 -NDKTPWKEFTQGLQFLPELGRDAIEIIGNDVQAEYNSLE---IMMRDFENKFW 161	
QY	173 KLKRLQAPGLPPSSALQQAALTILKIRFENVHDFIREIPQFL-ETVKTLLLPIYAQAN 231	
DB	162 -----SNRTANAIAVTAFSTVNTQIIRLKERFLAPENRPAFLNIYAQTAN 209	
QY	232 FHLNLLQQAELADEWNADHPISOIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKN 291	
DB	210 IDLILYQGSVYGKQWADINNRSTSP--PSSKDYQSLKGKIKDYNYCAETVYRSLTI 267	
QY	292 LRDEPNMKHSIFNDYRYMTITVLTISQFSLDIKRYRDSIGGIEYKGIKNELTREIYT 351	
DB	268 LKXKPHIQWDIVNRYRREALGALDLVALFPNYDICIYPTQ-----TRTELTRKVM 319	
QY	352 TEINPDRLPQLRVQPNLATWEYNLTRASPKLPSFLEQIFITYTENTNFGN--RLVGISNRD 409	
DB	320 PSFYLAQLORDIE-----TVENQLTHPP-SLFTWLNLNLYTTRERNPVLQVASLSGLQ 374	
QY	410 APT-----YSGNTTETLYGE-RTGSPPTTKTIRPFESYKSVITDRQSPVPVPIQPHFIINQ 464	
DB	375 ATSRYTQNTTISNPNQGVREGTPTKISLANYIYKLFMSQYRHPNDCLPISG---INE 431	
QY	465 IEL-----YLGSSNNTLYKYSAGSLSNYQNTTFFQPRKKDCNLVIDPGCSPPNPNVSHI 520	
DB	432 MSFYRSDYYGAGPAPVHYSGASPTNVIK-TYMNGPQ-----NALLISNDIS--INETSII 484	
QY	521 LSHFSLFTSYVIGLQILQIDTGV-LGWTHSSVDYRNAISDKLITMTPAKGNLDTNSK 579	
DB	485 LSDIKM-NYSRTGGV-YPLDYFGYSFAWTHTSVDDPNLIVPNRITQIPAVKASLTSPAR 542	
QY	580 VTGPGHTGNLVL-----SQQRLEITCTPNST---QSYFIRLRVATYGAGNLTLPNIS 632	
DB	543 VIVPGHTGDDLVALNSGTQSTQMTQCKTGFTGFSRQYGLRMRYAANSATV--SLS 600	
QY	633 LTPGVIGIPQRLNNTFSGTNY---NNLOYGDFGVFPFPS--TVTLPLNRNPIPIFNRA 687	
DB		
Db	601 YTLGQFRGTSFVIENTSRNNIIPDLYEYFKYKDYLOLQIITWLTPLANTITISMQQA 660	
Qy	688 -DVNSILIIDKIEPIPI 704	
Db	661 TGLLNQLIIDRIEFPYM 678	
RESULT 10		
CR8BA_BACUK		
ID	CR8BA_BACUK STANDARD; PRT; 1169 AA.	
AC	Q45705;	
DT	16-OCT-2001 (Rel. 40, Created)	
DT	16-OCT-2001 (Rel. 40, Last sequence update)	
DT	13-SEP-2005 (Rel. 48, Last annotation update)	
DE	Pesticidal crystal protein cry8Ba (insecticidal delta-endotoxin	
DE	CryVIIIb(a)) (Crystalline entomocidal protoxin) (134 kDa crystal	
DE	protein).	
GN	Name=cry8Ba; Synonyms=50C(b), cryVIIIb(a);	
OS	Bacillus thuringiensis subsp. kumamotoensis.	
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;	
OC	Bacillus cereus group.	
OX	NCBI_TaxID=132267;	
RN	[1]	
RP	NUCLEOTIDE SEQUENCE [GENOMIC DNA].	
RA	STRAIN=NRRL B-18746 / PS50C;	
RA	Michaels T.E., Fonceirada L., Narva K.E.;	
RT	"Process for controlling scarab pests with Bacillus thuringiensis	
RT	isolates.";	
RL	Patent number WO9315206, 05-AUG-1993.	
CC	-!- FUNCTION: Promotes colloidsmotic lysis by binding to the midgut.	
CC	epithelial cells of insects. Active on various scarabaeid beetles.	
CC	-!- DEVELOPMENTAL STAGE: The crystal protein is produced during	
CC	sporulation and is accumulated both as an inclusion and as part of	
CC	the spore coat.	
CC	-!- MISCELLANEOUS: Toxic segment of the protein is located in the N-	
CC	terminus.	
CC	-!- SIMILARITY: Belongs to the delta endotoxin family.	
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use as long as its content is in no way modified and this statement is not	
CC	removed.	
DR	EMBL; U04365; AAA21118.1; -; Genomic DNA.	
DR	HSP; P07130; 1DL.	
DR	InterPro; IPR001178; Endotoxin.	
DR	InterPro; IPR005638; endotoxin_C.	
DR	InterPro; IPR005639; endotoxin_N.	
DR	Pfam; PF03944; Endotoxin_C; 1.	
DR	Pfam; PF00555; Endotoxin_M; 1.	
DR	Pfam; PF03945; Endotoxin_N; 1.	
KW	Sporulation; Toxin.	
SQ	SEQUENCE 1169 AA; 133544 MW; 22BFCF5BDB699909 CRC64;	
	Query Match 20.0%; Score 775.5; DB 1; Length 1169;	
	Best Local Similarity 28.5%; Pred. No. 3.3e-41;	
	Matches 217; Conservative 136; Mismatches 289; Indels 119; Gaps 27;	
Qy	5 NDNNEYIIDSHTSPYFPRNSNDS-RYPYTNPNQPLQNTYKRWLNMCQN-TOYGDN 62	
Db	4 NNQVEYIIDATPS-----TSVNSNDRYPFANEPTNALQNNMDYKDYLRMSAGNVEYPS 59	
Qy	63 FEFPASADTTAAVAGTIVSGTLLAGIGLTSISGPIGIGIAIISFGTLITVFPAGQ 122	
Db	60 PEVFLSQD--AVKAAIDIVGKLTGLG-----VPFVGPIVSLYTLQIDILMPS--K 107	
Qy	123 DKTWTQTFIKMGEIFVDTPLTESIKQLKLTLEGFRILQSYNTALDDWRKLQAPGL 182	
Db	108 QKSQWELFMQVEELINQKIAEYARNKALSLEGLGNNYQLYLTALKEWKE----- 158	
Qy	183 PSSALQQAALTILKIRFENVHDFIREIPGFOLEYTKTLTLLPIYAAANFHLNLLQQGAE 242	

Db 159 NPGS--RALRDVNRPEILDSLFTQYMPSTRVNFVPELFTYTMANLHLLLRDASI 216
Qy 243 LADEWADIPHSQIEPNAGTSDDYKLLKENIPKYSNCANTYRTGLKNLRDSEPNMKWSI 302
Db 217 FGEWGL-----STSTINNYNRQMKLTAEYSDHCWKVETGLAKLGSSAKQWID 267
Qy 303 FNDYRYWITVLDTSQSFLYDIKRYDSIGIEVKGIKNELTREIYTEINFDRLPOL 362
Db 268 YNQFREMLTLLVDVVALFSNYDTRY-----PLATTAQLTREVTYDPLGAVDVPNI 319
Qy 363 RVQPNLATMEYNLTRASFK---LFSLEQPIFYTENTNF-----GNRLVGISNRD 409
Db 320 GSWYKAPSEFSELEKAAIRPPHVFDTYGLTVYTKRSFTSDRYMYWAGHQ---ISYKH 376
Qy 410 APTYSNTITETALYGERGTGPTTKTIIPFESYKV-----SIVTDQSPPVVS-----PI 456
Db 377 IGT-SSTFQ-MYGTQNLQSTSNF-DFTNYDIYKTLNCAVLLDIVPGYTYTFMGPE 433
Qy 457 QPHFIINQIELYNGSSNNTLKYSAGSLSNYQNTFFQPPRKOCNVLVIDPGCS--PNF 514
Db 434 TBFPMVNL-----NNTRK-----TLTYKPASKDIIDTRDSELELPPTSGQPNY 479
Qy 515 NNYSHLISH-----FSLFTSYVYVIGLQLOLDTGVLGTWTHSSVDRYNAISDKIITMIPA 570
Db 480 ESYSHRLGHITFYSSSTSYV-----PVFSWTHRSADLTNTVKSGETIQPGGK 529
Qy 571 GNNLDNNSKVGEPGHTGNNLVYLSQ-QRLTEITCETPNSTQSYFTRLRYATNAGNLTLP 629
Db 530 STIGRNTYIKRGYTGDDVALTRIGSCFQMIFPES-QRFRIRIRYASNETS-----584
Qy 630 NISLTIPGVIGIPPPORLNTFTSGTNNYNNLQYDFGYPQFPSTVTLPLNRRNIPFIFNRADV 689
Db 585 --VISLYGLNQSTLKFNTQYSKNENDLTYNDFKVIYPRVSNASSNIQRLSIGIQ 642
Qy 690 SNSIILDKIEPIPTSSMHQNRKQKLETIQTKINTFFTN 730
Db 643 NTNLFILDRIFIPVDTEYAEATD---LEAKKAVNALEFTN 680

RESULT 11
CR4BA_BACTI STANDARD; PRT; 1136 AA.
AC P05519; P11782; P16479;
DT 01-NOV-1988 (Rel. 09, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Pesticidal crystal protein cry4Ba (insecticidal delta-endotoxin
DE CryIVB(a)) (Crystalline entomocidal protoxin) (128 kDa crystal
DE protein).
GN Name=cry4Ba; Synonyms=bt8, cryD2, cryIVB(a), isrH3;
OS Bacillus thuringiensis subsp. israelensis.
OC Plasmid 72 Kb.
OG Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=1430;
RN [1]_TaxID=1430;
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=88157738; PubMed=2831510;
RA Tungpradabkul S., Sattasatian C., Panyim S.;
RT "The complete nucleotide sequence of a 130 kDa mosquito-larvicidal
RT delta-endotoxin gene of Bacillus thuringiensis var. israelensis";
RL Nucleic Acids Res. 16:1637-1638(1988).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=88185334; PubMed=2833395;
RA Chungjatupornchai W., Hofste H., Seurinck J., Angsuthanasombat C.,
RA Vaek M.;
RT "Common features of Bacillus thuringiensis toxins specific for Diptera
RT and Lepidoptera";
RL Eur. J. Biochem. 173:9-16(1988).
RN [3]

RP NUCLEOTIDE SEQUENCE.
RA Sen K., Honda G., Koyama N., Nishida M., Neki A., Sakai H., Himeno M.,
RA Komano T.;
RT "Cloning and nucleotide sequences of the two 130 kDa insecticidal
RT protein genes of Bacillus thuringiensis var. israelensis";
RL Agric. Biol. Chem. 52:873-878(1988).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=88329719; PubMed=2901387; DOI=10.1016/0378-1119(88)90229-6;
RA Yamamoto T., Watkinson I.A., Kim L., Sage M.V., Stratton R.,
RA Akande N., Li Y., Ma D.-P., Roe B.A.;
RT "Nucleotide sequence of the gene coding for a 130-kDa mosquitoicidal
RT protein of Bacillus thuringiensis israelensis";
RL Gene 66:107-120(1988).
RN [5]
RP NUCLEOTIDE SEQUENCE OF 1-134.
RX MEDLINE=88038331; PubMed=2890080; DOI=10.1007/BF00328128;
RA Angsuthanasombat C., Chungjatupornchai W., Kertbundit S.,
RA Luxananil P., Sattasatian C., Wilairat P., Panyim S.;
RT "Cloning and expression of 130-kd mosquito-larvicidal delta-endotoxin
RT gene of Bacillus thuringiensis var. israelensis in Escherichia coli";
RL Mol. Gen. Genet. 208:384-389(1987).
CC -I- FUNCTION: Promotes colloidsmotic lysis by binding to the midgut
CC epithelial cells of mosquitos.
CC -I- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -I- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -I- SIMILARITY: Belongs to the delta endotoxin family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; X07082; CAA30114.1; -; Genomic DNA.
DR EMBL; X05692; CAA29174.1; -; Genomic DNA.
DR EMBL; M00247; BAA00178.1; -; Genomic DNA.
DR EMBL; M02442; AAA22337.1; -; Genomic DNA.
DR EMBL; X07423; CAA30312.1; -; Genomic DNA.
DR PIR; S00398; USBS81.
DR PDB; 1W99; X-ray; A=84-641.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR Pfam; PF03944; Endotoxin_C; 1.
DR Pfam; PF00555; Endotoxin_M; 1.
DR Pfam; PF03945; Endotoxin_N; 1.
KW 3D-structure; Plasmid; Sporulation; Toxin.
FT CONFLICT 51 51 V -> D (in Ref. 4).
FT CONFLICT 65 65 T -> S (in Ref. 4).
FT CONFLICT 193 201 LINAQESL -> PHKTRMYV (in Ref. 4).
FT CONFLICT 203 204 RS -> C (in Ref. 3).
FT CONFLICT 205 206 AG -> C (in Ref. 4).
FT CONFLICT 205 205 A -> R (in Ref. 1).
FT CONFLICT 272 272 Y -> L (in Ref. 3).
FT CONFLICT 325 325 D -> Y (in Ref. 3).
FT CONFLICT 364 401 FGSNLTHQIQNSNVKTSITDTSPPSNRVTKMDPYKI ->
FT LVQYLKFLNLLIVLIVKLSQILAPPLIELOKWISTKF
FT (in Ref. 4).
FT CONFLICT 467 467 K -> N (in Ref. 4).
FT CONFLICT 496 496 K -> R (in Ref. 4).
FT CONFLICT 519 519 G -> GG (in Ref. 4).
FT CONFLICT 551 572 LNVSYLQVSGRGTITESTF -> IECDHMYKEFLEEQ
FT RLQVNYV (in Ref. 4).
FT CONFLICT 594 594 F -> N (in Ref. 4).
FT CONFLICT 687 690 ELYP -> GIIS (in Ref. 4).
FT CONFLICT 721 721 A -> R (in Ref. 4).
FT CONFLICT 823 832 SNRCETSAVP -> LIIVSVRECA (in Ref. 4).
FT CONFLICT 836 837 GN -> WD (in Ref. 4).
FT CONFLICT 902 902 E -> R (in Ref. 4).


```
FT CONFLICT 1015 1015 G -> V (in Ref. 4).
SQ SEQUENCE 1136 AA; 127764 MW; 8AC4E8C26FE3E9B5 CRC64;

Query Match
Best Local Similarity 19.6%; Score 758.5; DB 1; Length 1136;
Matches 232; Conservative 123; Mismatches 266; Indels 125; Gaps 33;

QY 28 DSRYPYNNPQPLQNTNYKEWLNMCQNTQYGDNFETFASADTIAAVSAGTIYSGTLA 87
DB 2 NSGYPLANDLQGSMMKNTNYKDWLACENNQQYGVN-----PAINSSSVSTALKVAGAI-- 55
QY 88 GIGGLTSISGPIGIIIGAIISFGTLITVFWPAG-EQDKTVMTOFIKMGIEIFVDTPLETESI 146
DB 56 ----LKFVNPAGTVLTVL---SAVLPLWPTNTPTPERVWDFMTNGLIDQTVTAYV 108
QY 147 KQLKLTLEGFRQILQSYNTALDDWRKLRLOAPGLPSSALQQAALTILKIRFNVHNDP 206
DB 109 RTDANAKMTVVKDYLDQYTKFTWKR-----EPNNQSYRTAVIT---QF-NLTSK 156
QY 207 IREIPGP-OLETVKTLPIYAQAANFHLNLQGAELADEWNADIHPSQIEPNAGTSD 264
DB 157 LRTAVFSLVGVYELLPIYAQVANFNLLIRDLGLNAQEWS-----LARSAG--D 207
QY 265 DYYKLLKENIPKYSNCAVYRTGLKRLDRBNPKWKSIFNDYRYMTITVLDTISQFSLY 324
DB 208 QLYNTWQYTKETIAHSITWYKGLDVLNRKSNQWITFNDYKREMTIOVLDIILFASY 267
QY 325 DIKY-RDSGGIEVGKIKNELTREIYTTINFDRLPQLRVQPNLATMEYNLTRASFCLP 383
DB 268 DPRYPADKIDNTKLS--KTEFTREIYTA-----LVESPSSKSIAALEAALTR-DVHLF 318
QY 384 SFLEQFIYFNTNFGNRLVGSINRDAPYNTIT---ETLYGER-TGSPTTKTIIRPPES 439
DB 319 TWLKRVDFT-NTIYQDLRFLSANKIGFSYNTSSAMQESGIYSGSGFNLTHQIQLNSN 377
QY 440 -YKVSIVTDROSPPVSPQIHPFIINQIELYNGSSNTLK---YSAGSLSNYQN----- 490
DB 378 VYKTSI-TDTSSP-----SNRVTKMDFYKIDGTLASYNSTIPT 415
QY 491 -----TTFOPPRKOCNLVIDPGSPNFNYSIHLSHFSLFYVYIGLQQLDITG-- 543
DB 416 PEGRLTTFGFSNEN-----TPN-QPTVNDYTHLSY-----IKTDVIDNSN 458
QY 544 --VLGTHSSVDRYNAISDKIITMIPAIGKNLDTNSKVIKPGHGTGGLNLYLOS-----Q 597
DB 459 RVSPAWTHKIVDPNNQIYTDAITQVPAVKSFNFLNATAKVIKPGHGTGGLVALTSNGTLS 518
QY 598 GRLEITCEIP--NSTQSYFIRLRYATNGAGTLPNLSLITPGVIGIPQRLNNTFSGTN 654
DB 519 GRMEIQKTSIFNDPSTRSYGLRIRYAANSP--IVLNVSIVLQGVSRGTTISTESTFSRPN 576
QY 655 Y---NNLQYDGFYQ-FPSTVTLPLNRN--IPFIFNRADV-SNSILIIDKIEFIPITSS 707
DB 577 NIPTDLKYEFRYKDPFDAIVPMRLSSNQLITIAIQLNMTSNQVIIDRIEIPITQS 636
QY 708 MHQNRKQKLETTQTKINTFTNHTK 733
DB 637 VLDETENQLESEREVNALTNDK 662
```

RESULT 12

```
Q7AL72 BACTI
ID Q7AL72 BACTI PRELIMINARY; PRT; 1136 AA.
AC Q7AL72;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, last sequence update)
DE Pesticidal crystal protein cry4BA.
GN Names=cry4BA; Synonyms=Bt8, cryD2, cryIVB(A), isrH3, pBt038;
OS Bacillus thuringiensis (subsp. israelensis).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=1430;
```

RESULT 13

```
Q5W7N9 BACTU
ID Q5W7N9 BACTU PRELIMINARY; PRT; 688 AA.
AC Q5W7N9;
```

```
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=22235415; PubMed=12324359;
RA DOI=10.1128/AEM.68.10.5082-5095.2002;
RA Berry C., O'Neil S., Ben-Dov E., Jones A.F., Murphy L., Quail M.A.,
RA Harris D., Zaritsky A., Parkhill J.,
RT "Complete sequence and organisation of pToxis, the toxin-coding
RT plasmid of Bacillus thuringiensis subsp. israelensis.";
RL Appl. Environ. Microbiol. 68:5082-5095(2002).
DR EMBL; AL731825; CAD30095.1; -; Genomic DNA.
SQ SEQUENCE 1136 AA; 127763 MW; 8AC4E8C26FE3E9B5 CRC64;
```

Query Match 19.6%; Score 758.5; DB 2; Length 1136;
Best Local Similarity 31.1%; Pred. No. 4e-40;
Matches 232; Conservative 123; Mismatches 266; Indels 125; Gaps 33;

```
QY 28 DSRYPYNNPQPLQNTNYKEWLNMCQNTQYGDNFETFASADTIAAVSAGTIYSGTLA 87
DB 2 NSGYPLANDLQGSMMKNTNYKDWLACENNQQYGVN-----PAINSSSVSTALKVAGAI-- 55
QY 88 GIGGLTSISGPIGIIIGAIISFGTLITVFWPAG-EQDKTVMTOFIKMGIEIFVDTPLETESI 146
DB 56 ----LKFVNPAGTVLTVL---SAVLPLWPTNTPTPERVWDFMTNGLIDQTVTAYV 108
QY 147 KQLKLTLEGFRQILQSYNTALDDWRKLRLOAPGLPSSALQQAALTILKIRFNVHNDP 206
DB 109 RTDANAKMTVVKDYLDQYTKFTWKR-----EPNNQSYRTAVIT---QF-NLTSK 156
QY 207 IREIPGP-OLETVKTLPIYAQAANFHLNLQGAELADEWNADIHPSQIEPNAGTSD 264
DB 157 LRTAVFSLVGVYELLPIYAQVANFNLLIRDLGLNAQEWS-----LARSAG--D 207
QY 265 DYYKLLKENIPKYSNCAVYRTGLKRLDRBNPKWKSIFNDYRYMTITVLDTISQFSLY 324
DB 208 QLYNTWQYTKETIAHSITWYKGLDVLNRKSNQWITFNDYKREMTIOVLDIILFASY 267
QY 325 DIKY-RDSGGIEVGKIKNELTREIYTTINFDRLPQLRVQPNLATMEYNLTRASFCLP 383
DB 268 DPRYPADKIDNTKLS--KTEFTREIYTA-----LVESPSSKSIAALEAALTR-DVHLF 318
QY 384 SFLEQFIYFNTNFGNRLVGSINRDAPYNTIT---ETLYGER-TGSPTTKTIIRPPES 439
DB 319 TWLKRVDFT-NTIYQDLRFLSANKIGFSYNTSSAMQESGIYSGSGFNLTHQIQLNSN 377
QY 440 -YKVSIVTDROSPPVSPQIHPFIINQIELYNGSSNTLK---YSAGSLSNYQN----- 490
DB 378 VYKTSI-TDTSSP-----SNRVTKMDFYKIDGTLASYNSTIPT 415
QY 491 -----TTFOPPRKOCNLVIDPGSPNFNYSIHLSHFSLFYVYIGLQQLDITG-- 543
DB 416 PEGRLTTFGFSNEN-----TPN-QPTVNDYTHLSY-----IKTDVIDNSN 458
QY 544 --VLGTHSSVDRYNAISDKIITMIPAIGKNLDTNSKVIKPGHGTGGLNLYLOS-----Q 597
DB 459 RVSPAWTHKIVDPNNQIYTDAITQVPAVKSFNFLNATAKVIKPGHGTGGLVALTSNGTLS 518
QY 598 GRLEITCEIP--NSTQSYFIRLRYATNGAGTLPNLSLITPGVIGIPQRLNNTFSGTN 654
DB 519 GRMEIQKTSIFNDPSTRSYGLRIRYAANSP--IVLNVSIVLQGVSRGTTISTESTFSRPN 576
QY 655 Y---NNLQYDGFYQ-FPSTVTLPLNRN--IPFIFNRADV-SNSILIIDKIEFIPITSS 707
DB 577 NIPTDLKYEFRYKDPFDAIVPMRLSSNQLITIAIQLNMTSNQVIIDRIEIPITQS 636
QY 708 MHQNRKQKLETTQTKINTFTNHTK 733
DB 637 VLDETENQLESEREVNALTNDK 662
```

```

DT 01-FEB-2005 (T-EMBLrel. 29, Created)
DT 01-FEB-2005 (T-EMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (T-EMBLrel. 29, Last annotation update)
DE Cry30-like.
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=1428;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=96-OK-85-24;
RA Ohgushi A., Wasano N., Saitoh H., Ohba M.;
RT "Identification and characterization of novel cry genes from an
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
EMBL; AB193814; BAD67157.1; -; Genomic DNA.
DR GO; GO:016787; F:hydrolyase activity; IEA.
DR GO; GO:005102; F:receptor binding; IEA.
DR GO; GO:0003952; P:defense response; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR GO; GO:0030435; P:sporulation; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR Pfam; PF03944; Endotoxin_C; 1.
DR Pfam; PF00555; Endotoxin_M; 1.
DR Pfam; PF03945; Endotoxin_N; 1.
SQ SEQUENCE 688 AA; 77438 MW; D5F1C265AEC6BD1 CRC64;

Query Match 19.0%; Score 734; DB 2; Length 688;
Best Local Similarity 31.2%; Pred. No. 7.7e-39;
Matches 238; Conservative 111; Mismatches 271; Indels 142; Gaps 34;

QY 1 MNQNDNNEVEIIDSHTSPYFPPNRNDS-----RYPVTNNPNQPLQNTYKELNWCQGN 56
DB 1 MDLYGNKNDSEILNASSN-----NSNNSTYPKYPLAHSRQDSMMENMYKEWINQC--- 51

QY 57 TOYGNFETFAF-ADT-IAAVSAGTVSGTLGIGGLTSISGPIGIGLTSISGFTGLIT 114
DB 52 ----ETINTFCTPIDIDNSVAATIGAVGAILA-----LIPGGEAIGFVLSTFTSLIP 101

QY 115 VFWPAGEQKTVWTFQIKWG-ELFVDTPLTESIK-----QLKQLTEGPRQILQSYNTA 167
DB 102 YLWPS---DTKIKWGDTKQGLQFLPELGNDAIEIIGNDVQSEYNSLKTFF---MQNFEDS 156

QY 168 LDDWRKLKELQAPGLPPSALQQAALTKIRPENVHDFIREIPGQLE-TYKTLPLPIY 226
DB 157 FTDWKKYRN-----RATAVAVTNDFSVRDQIIRLKDRLINPENKPAFLILY 204

QY 227 AQAAFHLLNLLQGAELADEWNADIHPSQTEPNAGTSDDYKLLKENIPKYSNYCANTYR 286
DB 205 AQTANFDLILYQRGALYADEWENDINRS-ISPILG-SKDYIISLAAKIKETNYCAETVR 262

QY 287 TGLKNLRDSENKWSIFNDYRYMTITVLDTSIQFSLYDIKRYRDSIGGIEVGKINKELT 346
DB 263 NSLNILKNKNTISWGTYNKYRREVTLCALDLAALFPNYDICIPIQ-----TKTELT 314

QY 347 RIYYTEINFDRLPQLRVQPNLATMEYNLTRASFKLPSFLEQEIFYTFENTNFGNRLVGIS 406
DB 315 RKVYMPSPFG---LQOSNYFQSLEGNALTHPP-SLFTMLNELNLYTIRENFNPALLVSS 370

QY 407 NRDAPTYF-----NTITELYGERTGSPPTKIRPFESYKVSIVTDQSPVPSPQPHF 460
DB 371 LSGLOAISRYTQNPWRISNPAQVRNGTPTQIGNNLFVYKLSU-----SOYHHPNECYF 425

QY 461 I---INQIELY-----LNGSSNNLTKYSAGSLSNQNTTFFQFPRKDCNLVIDPCGSPNFN 515
DB 426 IAGISDWTYKSDYNGNAPTQYQAGRSNNFIN-TFWNGPOE-----ASSSN 473

QY 516 NYS-----HILSHFSLF-----TY-SYVIGLQQLDITGVLTGWTSSVDRYNAISDKII 563
DB 474 NISIKQTNHILSDIKMYRGTGTGTPSYDFGYS-----FAWTHTSVNPDLNLPVNR 525

QY 564 TMIPAIGNNLDNTNSKVIIEGPGHTGGMNLVYL-----QSQGRLEITCETPNST---OSYFIR 616
DB 526 TQIPAVKADYLTSPAKYIAGPGHTGGDLVALLNAATQAGRMQIQCKTSGFTGASRRYGR 585

QY 617 LRYATNGA-----GNTLPNISITIPGVIGIPQRLNNTFSGTNNYNNLQYGDRCY 665
DB 586 IRYAANALTVSLSYTVQGGNTMTTFTTERTFL-----RPNNTIP-----TDLKYEEPKY 636

QY 666 FQFPS--TVTLPNLRNIPFIFNRADV-SNSILIIDKIEFFIPI 704
DB 637 KEYNQIITMAPQNTIVTIALQQLNAPFNQDLIIDRIEFFEYPM 678

RESULT 14
C19BA_BACUH
ID C19BA_BACUH STANDARD; PRT; 682 AA.
AC O86170;
DT 18-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Pesticidal crystal protein cry19Ba (Insecticidal delta-endotoxin
DE CryXIB(a)) (Crystalline entomocidal protoxin) (78 kDa crystal
DE protein).
DE Name=cry19Ba; Synonyms=cryXIB(a);
DE Bacillus thuringiensis subsp. higo.
OS Bacillus cereus group.
OG Plasmid.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=132266;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=98369730; PubMed=9704107;
RA Wang S.H., Saitoh H., Mizuki E., Higuchi K., Ohba M.;
RT "A novel class of mosquitocidal delta-endotoxin, Cry19B, encoded by a
RT Bacillus thuringiensis serovar higo gene."
RL Syst. Appl. Microbiol. 21:179-184(1998).
CC -!- FUNCTION: Promotes colloid-osmotic lysis by binding to the midgut
CC epithelial cells of mosquitoes.
CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; D88381; BAA32397.1; -; Genomic_DNA.
CC HSP; Q06117; J16.
CC InterPro; IPR001178; Endotoxin.
CC InterPro; IPR005638; endotoxin_C.
CC InterPro; IPR005639; endotoxin_N.
CC Pfam; PF03944; Endotoxin_C; 1.
CC Pfam; PF00555; Endotoxin_M; 1.
CC Pfam; PF03945; Endotoxin_N; 1.
CC Plasmid; Sporulation; Toxin.
SQ SEQUENCE 682 AA; 78491 MW; 5351EA63E2B042F7 CRC64;

Query Match 18.3%; Score 709; DB 1; Length 682;
Best Local Similarity 30.3%; Pred. No. 3.2e-37;
Matches 232; Conservative 112; Mismatches 277; Indels 144; Gaps 30;

QY 1 MNQNDNNEVEIIDS-----HTSPYFPPNRNDSRYPTNNPNQPLQNTYKELNWCQGN 55
DB 1 MNSYQNKVEIILDAKENTCHMSNCP-----KYPLANDPQMYLRNTHYKDWLNWCEE 53

QY 56 NTOYGNFETFAADTTAAVAGSPTVSGTLGAGLTSGTISGPIGIGLTAIIISFTGLTV 115
DB 54 -----ASYASSGSPQLFKVG-----GSIVAKILGMIPEVGPL-----LSWMVSL 92

```

```

QY 116 FWPAGEQDKTVMTQFIKMGEIFVDTPLTETSIKQKQLQTLQEGFRQILOSYNTALDDWFKLK 175
Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 93 FWTIEBKNTVWEDMKYVANLKKQELNTDLNRATSNLGLNESLNIYNRAAIAWQNK 152
Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 176 RLQAPGLPSSALQAAALTKIRFENVHNDFIREIPG-FOLETYKTLTLLPIYAQAANFHL 234
Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 235 NLPQQ-----GAELEADENADHPSQIEPNAGTSDDYKLLKENIPKSYNCANTYRTGLK 290
Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 202 LLLRDVAIYKELG-----YEST-----DVEFYNEQKYATEKYSNCVNTYKSGLE 248
QY 291 NLRDEPNMKSIFENDYRRYMTITVLTISQFSYDIKRYEDSIGGIEVKGKIKELTEIY 350
Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 249 S---KKQIGWSDNRRREMTSLVDLVALFPLYDTGLSPKDGKHVKA---ELTREIY 302
QY 351 TTEINFDRPLQRLRVQNLA--TMEYNLTASFKLFSPLEQFIFVTENTFNGRLVGLSNR 408
Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 303 SDVIN-DHYVGLMV-PYISFEHAESLYTRP-HAFTWLKGFRTVNSINSWTFLSGGENR 359
QY 409 DAPTY-SNTI-----TETLYGERTGSPPTKTIRPESYKVSIVTDROQPPVPSPHPH 460
Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 360 YELTHCEGTIYNGPFLQDTEYG-----GTSSVIDISNNSSIYNLWTKNWEIYPTDPV 414
QY 461 INQIELYNGSNNTLKYSAGGSLSNYQNTTFFQPPRKKDCNLVIDDPGCSN-PNFNYSH 519
Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 415 NITKINFSTDNSSESIYGAERMKPTVRTDFNF-----LLNAGNGPTTYNDYNH 467
QY 520 ILSHPSL-----FTSYVIGLQLQILDGTGLVGTWTHSSVDRYNAISDKIITMIPAIK 570
Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 468 ILSYMLINGETFGQKHGYSF-----APTHSSVDRYNTIVDPKIVQIPAVK 513
QY 571 GNNLDTSNKVIEGPGHTGNNLVYLSQSGRLTCTETPNSTQSYFIRLRVATNGAGNLPN 630
Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 514 TNLVGN--LIKPGHTGGDLLKLEVERFLSLRIKL-IASMTFRIRIRVYASNSGQMMIN 570
QY 631 ISLTIPGVGIPQRLNFTSGNYNLYGVDFG-----YFQFPSTVTLPLNRNIPFI 683
Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 571 IGYQNTYENIPT-----TSRDYTELKPEDQLVDTSYIYGGFS-----ISSNTLWL 619
QY 684 FNRADVSNLIIIDKIEFTPIITSSMQNREKQKLEIQTINTFF 728
Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 620 DN---FSGNPFVILDKIEFPLGTLNQ---AQQYDVTQNWANGMY 658

```

RESULT 15

Q6BE06 BACTU

ID Q6BE06 BACTU PRELIMINARY; PRT; 829 AA.

AC Q6BE06;

DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Cancer cell-killing Cry protein.

GN Name=cr41Ab1;

OS Bacillus thuringiensis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;

OC Bacillus cereus group.

OX NCBI_TaxID=1428;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=A1462;

RA Yamahita S., Saitoh H., Katayama H., Akao T., Mizuki E., Park Y.,

RA Ohba M., Ito A.;

RT "cell-killing toxin gene and other genes in 6,698bp DNA from Bacillus

thuringiensis";

RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB116651; BAD35163.1; -; Genomic DNA.

DR GO; GO:0005102; F:receptor binding; IEA.

DR GO; GO:0006952; P:defense response; IEA.

DR GO; GO:0009405; P:pathogenesis; IEA.

DR InterPro; IPR001178; Endotoxin.

DR InterPro; IPR005638; endotoxin_C.

```

DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF03944; Endotoxin_C_1.
DR Pfam; PF00555; Endotoxin_M; 1.
DR Pfam; PF03945; Endotoxin_N; 1.
DR Pfam; PF00652; Ricin_B_lectin; 1.
DR SMART; SM00458; RICIN; 1.
DR PROSITE; PS02031; RICIN_B_LECTIN; 1.
SQ SEQUENCE 829 AA; 93841 MW; EBFIF3BEC05F97A3 CRC64;

```

Query Match

Best Local Similarity 18.2%; Score 705; DB 2; Length 829;

Matches 213; Conservative 113; Mismatches 276; Indels 180; Gaps 28;

QY 1 MNQNDNNEVEIIDSHTSPYPNRSNDSRYPYNNPNQPLONTYKWLNNCOGN--TQ 58

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 1 MNQNDNNEVEIIDSHTSPYPNRSNDSRYPYNNPNQPLONTYKWLNNCOGN--TQ 58

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 59 YGDNFETFAADTTAAVSAGTIVSGTLLAGLGLTSIS-GPIGIIGAILIISFGTLITVFM 117

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 54 LGGGY-----SADVKDAVITSINIASYLL-----SVFPFPAAGVAGIL---GALLGLLW 99

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 118 PAGEQDKTVMTQFIKMGEIFVDTPLTETSIKQKQLQTLQEGFRQILOSYNTALDDWFKLKL 177

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 100 PTNTQ--AVWEAFNVTVEALLINQKLEVARSKAISELNGKLVLELYQDAADDWNE---- 153

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 178 QAPGLPSSALQAAALTKIRFENVHNDFIREIPGFOLETYKTLTLLPIYAQAANFHLNLL 237

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 154 ---NPGDLRNKNRVLt---EFRNVNGHFENSMSPFAVRNFEVNLPLPVYAAANLHLLLL 206

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 238 QOGNLADEWADHPSQIEPNAGTSDDYKLLKENIPKSYNCANTYRTGLKNLRD-EP 296

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 207 RDAVKFGEGMGSTDP-----GAERDDMYRLRSRTEIYDHCVTNTYNOGLQQAQSLQA 260

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 297 NM-----KWSIENDYRYMTITVLTISQFSLYDI 326

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 261 NVSDYSRYPTQYNQSGFSYREAKGEVGTENNWNLYNFAFRDMLILVDIIIAQPTTYD- 319

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 327 KRYRDSIGGIEVGKIKNELTREIYTTTEINFDRPLQRLVQPNLATMEYNLTR-----ASPKL 382

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 320 -----PGLYSRPVKSELTRVYTD---IRGTTWRSANLNTIDAENRMVGSRLQL 368

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 383 PSFLEQPIFYTENT-----NFGNRLVGISNRDAPTSYNTITETLYGBERTSGPTTKTIRP 436

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 369 FTWLTEMKFYIRNTGTSITSYTHGDLVGLKIRKTDNDQWLPLEGQNTS--YTRIDRP 426

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 437 -----FES--YKVSIVTQSPVSPVSIQPHFIINQIELYNGSSNNTL 477

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 427 GIBELKNYVYARTQQWFETRLQLQWNTDVLNAGTVGNEFWARDVPDYRNIVARSTR 486

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 478 KYSAGGSLSNYQNTTFFQFPKKDCNLVIDPGCSNPNFNPNYSHILSHFSFLTYSYVIGLQL 537

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 487 NH-----FIENHRLSWIKFEPVRDNC-----PFANPGYKQLSALL----- 521

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 538 QILDGTGLVGTSHSDRYNAISDKIITMIIPAIGNLNTDSKVIIEGPGHGTGNLVYLO-- 595

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 522 -----FGWTHNSVDPFNTIASDRIITQIPAVKGYLVNDGATVVRGPGTGDVRLPAY 574

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 596 -----SQGRLEITCETPNSTQSYFIRLYATNAGNTLPLNISLTIPGVIGIIPQ----- 644

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 575 NQOWTQLRVKVRPSTTARTAGYNNRIRYASEG-----NANLFVGKYDYDTRNFVETGNY 628

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 645 RLANTRSST-NYNNLOYGD-FGYFQFPSTVTLPLNRNIPFIFNRADVNSLIIIDKIEFI 702

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 629 AVNQTFSGSTYNSFKYDAIGAFAANEEEFRIELRCN-----SGGPIYIDKIEFI 678

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 703 PI 704

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 679 PV 680

Db ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

Search completed: December 4, 2005, 13:08:33

Job time : 234 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2005, 12:42:36 ; Search time 48 Seconds
(without alignments)
1265.971 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869

Sequence: 1 MNQNNNEVEIIDSHTSPY.....KLETIQKINTFTNHTKTL 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5.COMB.pep:*

2: /cgn2_6/ptodata/1/iaa/6.COMB.pep:*

3: /cgn2_6/ptodata/1/iaa/H.COMB.pep:*

4: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep:*

5: /cgn2_6/ptodata/1/iaa/RE.COMB.pep:*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1060.5	27.4	1180	2	US-09-224-024-28
2	1060.5	27.4	1180	4	PCT-US94-07902-28
3	895.5	23.1	686	2	US-09-224-024-31
4	895.5	23.1	686	4	PCT-US94-07902-31
5	775.5	20.0	1169	1	US-08-315-468-4
6	676	17.5	651	1	US-08-315-468-6
7	676	17.5	651	2	US-07-941-650A-4
8	676	17.5	659	2	US-08-996-441B-112
9	676	17.5	659	2	US-08-993-722A-112
10	676	17.5	659	2	US-08-993-170A-112
11	676	17.5	659	2	US-08-993-775B-112
12	676	17.5	659	2	US-09-427-770-112
13	676	17.5	659	2	US-09-427-769-112
14	663	17.1	652	2	US-08-996-441B-38
15	663	17.1	652	2	US-08-993-722A-38
16	663	17.1	652	2	US-08-993-170A-38
17	663	17.1	652	2	US-08-993-775B-38
18	663	17.1	652	2	US-09-427-770-38
19	663	17.1	652	2	US-09-427-769-38
20	660	17.1	652	2	US-08-996-441B-110
21	660	17.1	652	2	US-08-993-722A-110
22	660	17.1	652	2	US-08-993-170A-110
23	660	17.1	652	2	US-08-993-775B-110
24	660	17.1	652	2	US-09-377-466B-4
25	660	17.1	652	2	US-09-427-770-110
26	660	17.1	652	2	US-09-427-769-110
27	660	17.1	652	2	US-10-232-665-4

28 660 17.1 652 4 PCT-US92-00040-2 Sequence 2, Appli
29 658 17.0 652 2 US-08-996-441B-28 Sequence 28, Appl
30 658 17.0 652 2 US-08-996-441B-36 Sequence 36, Appl
31 658 17.0 652 2 US-08-993-722A-28 Sequence 28, Appl
32 658 17.0 652 2 US-08-993-722A-36 Sequence 36, Appl
33 658 17.0 652 2 US-08-993-170A-28 Sequence 28, Appl
34 658 17.0 652 2 US-08-993-170A-36 Sequence 36, Appl
35 658 17.0 652 2 US-08-993-775B-28 Sequence 28, Appl
36 658 17.0 652 2 US-08-993-775B-36 Sequence 36, Appl
37 658 17.0 652 2 US-09-427-770-28 Sequence 28, Appl
38 658 17.0 652 2 US-09-427-770-36 Sequence 36, Appl
39 658 17.0 652 2 US-09-427-769-28 Sequence 28, Appl
40 658 17.0 652 2 US-09-427-769-36 Sequence 36, Appl
41 656 17.0 652 2 US-09-377-466B-6 Sequence 6, Appl
42 656 17.0 652 2 US-10-232-665-6 Sequence 6, Appl
43 655 16.9 652 2 US-08-996-441B-18 Sequence 18, Appl
44 655 16.9 652 2 US-08-996-441B-60 Sequence 60, Appl
45 655 16.9 652 2 US-08-993-722A-18 Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-09-224-024-28
; Sequence 28, Application US/09224024
; Patent No. 6056953
; GENERAL INFORMATION:
; APPLICANT: Leslie Hickie
; APPLICANT: Jewel Payne
; TITLE OF INVENTION: Materials and Methods for the Control of
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/224,024
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/856,226
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA79
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1180 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-224-024-28

Query Match 27.4%; Score 1060.5; DB 2; Length 1180;
Best Local Similarity 35.1%; Pred No. 1.2e-89;
Matches 271; Conservative 117; Mismatches 279; Indels 105; Gaps 24;

Oy 1 MNQNNNEVEIIDSHTSPYFPNRSNDSRYPTNNPQNPLQNTYKWLNNCOGNTQYG 60

Db 1 MNPYQKNYETLNASQKL--NISNNYTRPIENSPKQLQSTNYKDWLNMCQOQYG 58
Qy 61 DNFTETASADTAAVAGSAGTIVSGTLLAGIGLTSISGPIGIIAIIISFGTLITVFWPAG 120
Db 59 GDFETPIDS--GELSAITIVGTVLTGFGFTT---PLGL---ALIGFGTLIPVLFPQA 108
Qy 121 EQDKTWTQFIKMGELFVDTPLTESIKQLKQLEGRQILQSYNTALDWRKLRLOAP 180
Db 109 DQSNLT--WSDFITQTKNIKEIASTYISNANKILNRSFNVISTYHNLKTWE-----NNP 162
Qy 181 GLPPSSALQQAALTAKIRFENVHNDIREIP--GFQLETYKTLPLPIYAAANFHLNLLQ 238
Db 163 NPQNTQDVRTQILQVHVHFQNVIPELVNSCPNPDSDYNNILVLSYAQAANLHLTVLN 222
Qy 239 QGAELADEWNADTHPSQIEPNAGTSDDYKLLKENIPKYSNCANTYRTKKNLRDBP-- 296
Db 223 QAVKFEAYLKNRRQFDYLEP-LPTAIDYYPVLTKAIEDYNYCVTYKGLNLKTTTDS 281
Qy 297 ----NMKWSIFNDYRYMTITVLDTSIQFSLYDIKRYRDSIGGIEVKGIKNELTREIYTT 352
Db 282 NLDGNINMNTYRTKMTTAVLDLVALFPNYDVGKY-----PIGVQSELTREIYQV 333
Qy 353 EINFDELQPLRVQPNLATMEYNLTRASFKLFSLEQEIFY-----TENTNFGNR----- 401
Db 334 -LNFEESPKYKY--DFQYQEDSLTRRP-HLFTWLDLSLNFYEKAQTPNPNFTSHYNNMFY 389
Qy 402 -LVGISNRDAPTYNTITETLYGERTGSPPTKTRPFESYKVSIVTDROQPPVSPIQPHF 460
Db 390 TLDNISQKSSVFGNHNVTDKL--KSLGLATNIYI-----PLGL---ALIGFGTLIPVLFPQA 108
Qy 461 IINQIEL---YLANGSSNNT-----LKYSAGGSLSNYQ-NTTFFQPRKDC 502
Db 423 LLNVISLDNKYLDYNNISKMDFFITNGTRLEKELTAGSGQITVDVKNKIFGLPKLR 482
Qy 503 NLVIDPGCSPNFNYSKSHLSHFSLFTYSYVIGLQILQDGTGVGTHSSVDRYNAISDKI 562
Db 483 ENQAIPTLFPTDYNYSHILSFILKSLIPATYKTVQY-----TPAWTHSSVDPKNTIYTHL 537
Qy 563 ITMIPAIKGNLDNTSKVIEGPGHTGCGNLVYLOSQGRLEITCETPNSTQSYFIRLYATN 622
Db 538 TTQIPAVKANSLGTASKVQVQPGHTGGDLI--DFKDHFKITCQHSNFQOQSYFIRIRYASN 595
Qy 623 GAGNTLPNLSLTPGIVIGIPPPORLNNTFSGTNNYNNLQYDGFQFPSTVTLPLNRNIPF 682
Db 596 GSANTRAVINLSIPGVAEL-GMALNFTFSGTDYTNLKYKDFQYLFESNEVKFAPQNISL 654
Qy 683 IPNRADV-SNSLIDIKIEPIPTSSMHQNRKQKLETTQTKNTFTNHTK 733
Db 655 VFNRSVYNTTTLVDIKIEBFLPITRSIREDRKQKLETVQOIIINTFYANPIK 706

RESULT 2

PCT-US94-07902-28
; Sequence 28, Application PC/TUS9407902
; GENERAL INFORMATION:
; APPLICANT: Street address: 4980 Carroll Canyon Road
; APPLICANT: City: San Diego
; APPLICANT: State/Province: California
; APPLICANT: Country: US
; APPLICANT: Postal code/Zip: 92121
; APPLICANT: Phone number: (619) 453-8030
; APPLICANT: Telex number: Fax number: (619) 453-6991
; TITLE OF INVENTION: Materials and Methods for the Control of
; TITLE OF INVENTION: Calliphoridae Pests
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA

; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07902
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA79
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1180 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US94-07902-28

Query Match 27.4%; Score 1060.5; DB 4; Length 1180;
Best Local Similarity 35.1%; Pred. No. 1.2e-89;
Matches 271; Conservative 117; Mismatches 279; Indels 105; Gaps 24;
Qy 1 MNQNDNNEYIIDSHTSPYFPNRSNDSRYPYTNPNQPLQNTNYKEWLNMCQNTQYG 60
Db 1 MNPYQKNYETLNASQKL--NISNNYTRPIENSPKQLQSTNYKDWLNMCQOQYG 58
Qy 61 DNFTETASADTAAVAGSAGTIVSGTLLAGIGLTSISGPIGIIAIIISFGTLITVFWPAG 120
Db 59 GDFETPIDS--GELSAITIVGTVLTGFGFTT---PLGL---ALIGFGTLIPVLFPQA 108
Qy 121 EQDKTWTQFIKMGELFVDTPLTESIKQLKQLEGRQILQSYNTALDWRKLRLOAP 180
Db 109 DQSNLT--WSDFITQTKNIKEIASTYISNANKILNRSFNVISTYHNLKTWE-----NNP 162
Qy 181 GLPPSSALQQAALTAKIRFENVHNDIREIP--GFQLETYKTLPLPIYAAANFHLNLLQ 238
Db 163 NPQNTQDVRTQILQVHVHFQNVIPELVNSCPNPDSDYNNILVLSYAQAANLHLTVLN 222
Qy 239 QGAELADEWNADTHPSQIEPNAGTSDDYKLLKENIPKYSNCANTYRTKKNLRDBP-- 296
Db 223 QAVKFEAYLKNRRQFDYLEP-LPTAIDYYPVLTKAIEDYNYCVTYKGLNLKTTTDS 281
Qy 297 ----NMKWSIFNDYRYMTITVLDTSIQFSLYDIKRYRDSIGGIEVKGIKNELTREIYTT 352
Db 282 NLDGNINMNTYRTKMTTAVLDLVALFPNYDVGKY-----PIGVQSELTREIYQV 333
Qy 353 EINFDELQPLRVQPNLATMEYNLTRASFKLFSLEQEIFY-----TENTNFGNR----- 401
Db 334 -LNFEESPKYKY--DFQYQEDSLTRRP-HLFTWLDLSLNFYEKAQTPNPNFTSHYNNMFY 389
Qy 402 -LVGISNRDAPTYNTITETLYGERTGSPPTKTRPFESYKVSIVTDROQPPVSPIQPHF 460
Db 390 TLDNISQKSSVFGNHNVTDKL--KSLGLATNIYI-----PLGL---ALIGFGTLIPVLFPQA 422
Qy 461 IINQIEL---YLANGSSNNT-----LKYSAGGSLSNYQ-NTTFFQPRKDC 502
Db 423 LLNVISLDNKYLDYNNISKMDFFITNGTRLEKELTAGSGQITVDVKNKIFGLPKLR 482
Qy 503 NLVIDPGCSPNFNYSKSHLSHFSLFTYSYVIGLQILQDGTGVGTHSSVDRYNAISDKI 562
Db 483 ENQAIPTLFPTDYNYSHILSFILKSLIPATYKTVQY-----TPAWTHSSVDPKNTIYTHL 537
Qy 563 ITMIPAIKGNLDNTSKVIEGPGHTGCGNLVYLOSQGRLEITCETPNSTQSYFIRLYATN 622
Db 538 TTQIPAVKANSLGTASKVQVQPGHTGGDLI--DFKDHFKITCQHSNFQOQSYFIRIRYASN 595


```
; REFERENCE/DOCKET NUMBER: MA79
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 686 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US94-07902-31

Query Match      23.1%; Score 895.5; DB 4; Length 686;
Best Local Similarity 31.9%; Pred. No. 1.4e-74;
Matches 245; Conservative 127; Mismatches 267; Indels 129; Gaps 30;

QY 1 MNQNNNEVEIIDSHTSPYFPNRRNSDRYPYTNPNQPLQNTNYKEWLNMCQNTQYG 60
DB 1 MNPYQKNYEIPNAPSGF--SKSNYSRYPLANKPNQPLKNTNYKDWLNVCODNQYQ 58
QY 61 DNFETPASADTTAAVAGATVSGTLLAGIGLTSISGPIGIIAGIISFGTLITVFWPAG 120
DB 59 NNAGNFVSSETIVGVAGIIVGTMLG-----APAAP--VLAAGIISFGTILLPIFW-QG 109
QY 121 EODKTWTOFIKNGEIFVDTPLTE---SIKQLKLTLEGFRQILQSYNTALDDWRKLR 177
DB 110 SDPANVQDLLNIG---GRPIQEIKNINVLTSIVTPIKNQLDKYQEFDFKWEPA-- 163
QY 178 QAPGLPSSALQQAALTAKIRFENVN---DFREIPGFOLEYTKYKTLILLBIYQAANFHL 234
DB 164 -----THANAKAVHDLFTTLEPIIDKOLDMLKNNASRYPT-----LPAYAQIATWHL 211
QY 235 NLLQQAELADENADIHPSQIBPNAGTSDYVK--LLKENIPKYSNYCANTYRTGLNLR 293
DB 212 NLLKHAATYNIW---LQNGQINPSTFNSNYIQQYLKRIQBYTDYCIOTYNAGLTMIR 268
QY 294 DBPNMKWSIFNDYRYMTITVLTISQFSLYDIKRYRDSIGGIEVKGKNELTREIYITE 353
DB 269 TTNATWNTMYNTRLEWTLTLDLIAIFPNYDEKY-----PIGVKSELREYV-TN 319
QY 354 INFDRLPQURVQNPLATMEYNLTRASFKLFSLFLEQIFITYENTN-----FGNRL 402
DB 320 VNSDTP-----RTITELENGLTR-NPTLFTWINGRFRYTRNSRDILDPYDIFSFTGNOM 372
QY 403 VGISNRDAPTSNTITELYGRTG--SPTTKTIRPFESY----KVISIVTRQSPVPSP 455
DB 373 -----AFTHNDDRNIIWGAHVGHIIISQDTSKVPFPYRNKPIDKVEIVRHREYSDII- 424
QY 456 IQPHFTIINOIELYNGSNNTLKYSGGSL--SNYQNTTFFQPRKXDCNLVIDPGGSPNF 514
DB 425 -----YEMIFFNSSEVFPYSSNTIENNYKRTDSYMIKQITWKN----- 464
QY 515 NNYSHILSHFSLFYTYVYIGLQILDTGVLGWTHSSVDRYNAISDKIITMIPAIGNNL 574
DB 465 KEYGHTLSVIKTDNYIFSVVRRRRV---AFSMTHTSDVPQNTIDLDNIQIHALKALKV 521
QY 575 DTNSKVIIEGPGTGNLVLOSQRLEITCETNSQSYFIRLRYATNGAGNTLPLISLT 634
DB 522 SSDSKIVKPGHGTGGDLVLKDSMDFRVRP-LKNVSRQYQVRIYATNA-----PKTTVF 575
QY 635 IPGVIGIPQRLNNTPSGTNYN--NLQYDGFYGFQFPSTV-----TLPNLRNIPFI 683
DB 576 LTGIDTISVE-LPSTTSRQNPNAITDYADFGVYFPTRTVPNTKFEGETLLMT-----L 629
QY 684 FNRADVSNILIDKIEPIPTSSMEHQNRKQKLETIQTKINTFFTNH 731
DB 630 YGTPNHSYNI-YIDKIEFIPITQSVLDYTEKQNIKTQKIVNDLFVNN 676

RESULT 5
US-08-315-468-4
; Sequence 4, Application US/08315468
```

```
; Patent No. 5554534
; GENERAL INFORMATION:
; APPLICANT: Michaels, Tracy Ellis
; APPLICANT: Foncerrada, Luis
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Process for Controlling Scarab Pests
; TITLE OF INVENTION: with Bacillus thuringiensis Isolates
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/315,468
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,941
; FILING DATE: 01 FEB 1993
; APPLICATION NUMBER: 07/828,430
; FILING DATE: 30-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/808,316
; FILING DATE: 16-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA73.C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1169 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: Kumamotoensis
; INDIVIDUAL ISOLATE: 50C
; IMMEDIATE SOURCE:
; LIBRARY: LambdaGEM-11 library of L. Foncerrada
; CLONE: 50C(b)
; US-08-315-468-4
```

```
Query Match      20.0%; Score 775.5; DB 1; Length 1169;
Best Local Similarity 28.5%; Pred. No. 6.4e-63;
Matches 217; Conservative 136; Mismatches 289; Indels 119; Gaps 27;

QY 5 NDNNEVEIIDSHTSPYFPNRRNSDRYPYTNPNQPLQNTNYKEWLNMCQNTQYGDN 62
DB 4 NNQNEVEIIDSHTSPYFPNRRNSDRYPYTNPNQPLQNTNYKEWLNMCQNTQYGDN 59
QY 63 FETTFASADTTAAVAGATVSGTLLAGIGLTSISGPIGIIAGIISFGTLITVFWPAG 122
DB 60 PEVFLSEQD--AVKAAIDIVGKLLTGLG-----VPFVGPIVSLYTQLDILWPS--K 107
QY 123 DKTWTOFIKNGEIFVDTPLTESIKQLKLTLEGFRQILQSYNTALDDWRKLRQAPGL 182
DB 108 QKSQWEIFMQVEELINOKIAEYARNKALSEGLEGNNGNYQLYLTAEWKE----- 158
```

183 PPSSALQOAAALTLKIRFENVHNDIFIREIPGFOLEYTKTLPLPIVAAANFHLNLLQGA 242
159 NPNGS--RALRDVNRREILDSLTQYPSRVNFVFPFTVTYMAANLHLLLRDASI 216
243 LADBNADIHPSQIEPNAGTDDYYKLLKENIPKYSNYCANTYRTGLKNLRDEPNMWSI 302
217 FGBEWGL-----STSTINNYNRQMKLTAEYSDHCVKWTETGLAKGSSAKQWID 267
303 FNDYRRYMTITVLDTISQFSLYDIKRYRDSGGIEVGKIKNELREIYTTINFDRLPOL 362
268 YNOFRREMTITVLVDVALFSNYDRTY-----PLATTAQLTREVVYDPLGAVDVENI 319
363 RVQPNLATMEYNLTRASF--LFSFLEQFIYTTENTNF-----GNRLVGSNRD 409
320 GSWYDKAPSESEIEKAIRPHVDYITGLTYTKRSFTSDRYMRVWAGHQ---ISYKH 376
410 APTYSNTITELYGERTGSPTKTIRPESYKV-----SIVTDROSPVVS-----PI 456
377 IGT--SSFTQ--MYGTNQLQSTSNF-DFTNYDIYKTLNSGAVLLDIVPGYTYTFFGMP 433
457 QPHFIINQIELYNGSSNNTLYKSAGGSLSNYQNTFFQFPKDKCNLVIDPGCS--PNF 514
434 TEFPMVNL-----NTRK-----TLTYKPSKDIIDTRDSELELPPETSGQPNY 479
515 NYVSHILSH-----PSLTYSYVIGLQILDGTGLGHTHSSVDRYNALSDKLIITMIPAIK 570
480 ESYSHRUGHITFYSSSTSTV-----PVFSWTHRSADLTNTVKSGBEITQIPGGK 529
571 GNNLDTNSKVIEGPGHTGNNLVYLSQ--GRLEITCETPNSTQSYFIRLRYATNGAGNLT 629
530 SSTIGRNTYIYKRGYTGGLDVALTDRIKGCSEFQWIPES--QRFIRIRVSNETS----- 584
630 NISLTIGVTGIPQRLNNTFSGYNNYNNLVQDFGFGYFQFPSTVTLPLNRNIPFIPNRAD 689
585 --VISYGLNQSGLTKFNQYNSKNENDLYNDPKYIEYPRVSNASSNTRLSIGT 642
690 SNSLITDKIEFTIPITSSMHQNRKQKLETTQTKINTFTN 730
643 NTNLFILDRIEFIPVDETYEATD---LEAKKAVNALFTN 680

RESULT 6
US-08-315-468-6
Sequence 6, Application US/08315468
Patent No. 5554534
GENERAL INFORMATION:
APPLICANT: Michaels, Tracy Ellis
APPLICANT: Focerrada, Luis
TITLE OF INVENTION: Process for Controlling Scarab Pests
TITLE OF INVENTION: with Bacillus thuringiensis Isolates
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/315,468
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/014,941
FILING DATE: 01 FEB 1993
APPLICATION NUMBER: 07/828,430
FILING DATE: 30-JAN-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/808,316
FILING DATE: 16-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA73.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
STRAIN: tolworthi
INDIVIDUAL ISOLATE: 43F
IMMEDIATE SOURCE:
CLONE: E. coli XL1-Blue (pML198-4), NRRL B-18291
FEATURE:
NAME/KEY: Protein
LOCATION: 1..651
US-08-315-468-6

Query Match 17.5%; Score 676; DB 1; Length 651;
Best Local Similarity 28.0%; Pred. No. 5e-54;
Matches 208; Conservative 116; Mismatches 286; Indels 134; Gaps 28;

QY 1 MNQNDNNEYIIDSHTSPYFNRNSDSRYPTVNNPQLONTYKELNMCQNTQYG 60
DB 1 MNPNN-RSEYDTIKVTPNSLP---TNNQYPLADNPSTLEELNYKEFLMTADNST-- 54
QY 61 DNFETASADTAAVSACTIVSGTLLAGIGLTSISGPIIGIIGAIISFGTLITVFPAG 120
DB 55 ---EVLDSSTVKDAVGTGIVVVGQILGVG-----VPPAGALTSTFYQSFLNFWPS- 102
QY 121 EODKTVTQPTKMGHIFVDTPLTESI KQLKLTLEGFRQILQSYNTALDDWRKLRLOAP 180
DB 103 --DADPKAFMAQVEVLIDKKIEYAKSKALAELOGLQNNPFDYVNALDSWK- 153
QY 181 GLPSSALQOAAALTLKIRFENVHNDIFIREIPGFOLEYTKTLPLPIVAAANFHLNLLQ 240
DB 154 -APVNLRSRRSQDRIRELFSQAESHFRNSMPSFAVSKFEVLFLPTVYQAANTHLLLLK 212
QY 241 AELADEMNADIHPSQIEPNAGTSD--YYKLLKENIPKYSNYCANTYRTGLKNLRDEPN 297
DB 213 QVGEWGY-----SSEDIAEFYQRLKLTQOYTDHCYVNWVYVGLNSLGSTY 260
QY 298 MKWSIFNDYRRYMTITVLDTISQFSLYDIKRYRDSIGGIEVGKIKNELREIYTTINFD 357
DB 261 DAWKFNRRFRREMTITVLVDLIVLFPFYDRLYS-----KGVKTELTRDITDPI-FT 311
QY 358 RLPQLRVQPNLATMEYNLTRASF--LFSFLEQFIYTTENTFNGNRL-VGINSRDAPTY--- 413
DB 312 LNALQYEGPTFFSISIENSIRKP--HLFDYLRIGTEFT-----RLRPGYSGKDSFNW 362
QY 414 -----SNTITELYGERTGSPTKTIRPPES---YKVSIVTDROSPVVSIPQH 459
DB 363 NYVETRPSIGSNDITISPFYGDKSIEPIQKL--SFDGQKVYRTIANTIDAAFPDGKI--Y 418
QY 460 FIINQIEL--YLANGSSNNTLYKSAGGSLSNYQNTFFQFPKDKCNLVIDPGCS--PNFN 516
DB 419 FGVTKVDFSQYDDQKNET-----STQYDSKRYNGYLGADSIDQLPPTTDEPLEKA 471
QY 517 YSHILSHFSLFTYSYVIGLQILDGTGLGHTHSSVDRYNALSDKLIITMIPAKGNLDT 576
DB 472 YSHQLNVAECFL-----MQDRRGITPFTTWTTHRSVDFENTIDAEKITQLPVVYKAYALS 525


```

; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996.441B
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: MECO:151
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-996-441B-112

Query Match 17.5%; Score 676; DB 2; Length 659;
Best Local Similarity 28.0%; Pred. No. 5,1e-54;
Matches 208; Conservative 116; Mismatches 286; Indels 134; Gaps 28

Qy 1 MNQNNDNNEYEIIISHTSYPYPNPNNSDRSRYPTNNPQIONTNYKXEWLNMCQNTQYG 60
Db 9 MNPNN-RSEYDTIKVTPNSLSP---TNNQYPLADNPNTLEELNYKFLRMTADNST-- 62
Qy 61 DNFTFPASADTIAVSAAGTIVSGTLLAGIGLTSISGPIGIGAILISFGTLITVFWPAG 120
Db 63 ---EVLDSSTVKDVAVTGSGVVGQTLGVVG-----VPPAGALTSPYQSFNLAINWPS- 110
Qy 121 EQDKTVMVTFQIKMGGEIFVDTPLTETSIKQLQTLQEGFQIQLQSVNTALDDWRKLRLOAP 180
Db 111 --DADPKWAFNAQVEVLIDKKIEYAKSKALAEQLQNNPFDYVALDSWKK----- 161
Qy 181 GLPSSALQQAALTKIRFENVHNDFTREIFGQLETKYTKLLLPITYAQAAPHMLNQOG 240
Db 162 -APVNLRSRRSQDRIRELFSQAESHFRNSMPFSFAVSKPEVLPLPTYAQAANTHLLLKDA 220
Qy 241 AELADENWADIHPQIEBNAGTSD--YYKLLKENIPKYSNYCANTYRTGLKNLRDPN 297
Db 221 QVFGEEGY-----SSEDAIEFYORQLKLTQYTDHCNVNWNVGLNSLRGSTY 268
Qy 298 MKWGIFFNDRYMYTITVLDTISQFSLYDIKKYRDSIGGIEVKGIKNELTREIYTTINPD 357
Db 269 DAWKFNRFREMTLVLDLILVFPFYDRLYS-----KGVTETLRDITFDPI-FT 319
Qy 358 RLPOLRVQPNLATMEYNLTRASFKLFSLEOFIFYTENTNFGNRL-VGISNRDAPTY--- 413
Db 320 LNALQYCGPTFSIENSIRKP-HLFDYLRGIEFHT-----RLRPGYSGKDSFNYWSG 370
Qy 414 -----SNWITETLYGERTGSPPTKIRPPES---YKVSIVTRQSPPVSPIDPH 459
Db 371 NYVETRFPSIGSNDRTTSPFYGDKGIEPIQKL--SPDGQKVYRTTANTDIAAPPDGI--Y 426
Qy 460 FIINQIEL-YLNGSSNNTLKYSGAGSLSNYQNTIFFQPRKKOCNVLDPGCS--PNFN 516
Db 427 FGVTKVDFSQVDDQKNET-----STQYDSKRYNGYLGQAQDSIDQUPPETTDEPLEKA 479
Qy 517 YSHLSHFSLPTYGVIGLOQLILDTVGLWTHSSVDRYNAISDKIIITWIPAIKGNNDLT 576

```

Db 63 ----EVLDSSTVDAVGTGIVGQILGWG-----VFPAGALTSFYQSFNLAIWPS- 110
Qy 121 EODKTVWTFIKMGEIFVDPPLTESIKQLKLTLEGPRLQLQSYNTALDDWRKLRLOAP 180
Db 111 --DADPKAFMAQVEVLIDDKIEYAKSKALAELOGLQNNFEDVYVALDSWK----- 161
Qy 181 GLPPSSALQQAULTKIRFENVHNDIFREIPGQLETYKTLTLLPIYAQAANFHLNLQOG 240
Db 162 -APVNLRSRQRIRELFSAQSHFRNSMPSPAVSKFEVLFLPTYAQAANTHLLLKDA 220
Qy 241 AELADEWNAIHPSQIEPNAGTSD--YYKLKENIPKYSNYCANTYRTGLKNLRDEPN 297
Db 221 QVFEGEWY-----SSEDIAEFYQRLKLTQYTDHCNVNWNVGLNSLRGSTY 268
Qy 298 MKWSIFNDYRRYMTITVLDTISQSLYDIKRYRDSIGGIEVKIKNELTRIETVTEINF 357
Db 269 DAWKFNRRFRREMTLVLDLIVLFPFYDVLYS-----KGVKELTRDIFTDPI-FT 319
Qy 358 RLPQLRVQPNLATMEYNLTRASPKLFSLEQFIYFTENTNFGNRL-VGISNRDAPTY--- 413
Db 320 LNALQEVGPTFSSIEHSIRKP--HLFDYLRGIEFHT-----RLRFGYSGKDSFNWWSG 370
Qy 414 -----SNTITETLYGERTSGPTTKTIRPFES---YKSVITVDRQSPVPSPIQPH 459
Db 371 NYVETRPSIGSNDTITSPFYGDKSIEPIQKL--SFDQKQVYRTIANTDIAAFPDKGI--Y 426
Qy 460 FIINQIEL-YLNGSSNNLTKYSAGGSLSNQTNTFFQPKKDCNCLVIDPGCS--PNFNN 516
Db 427 FGVTKVDFSQYDQKNET-----STQYDSKRYNGYLGQDSIDQLPETTDEPLEKA 479
Qy 517 YSHILSHFSLPTYVYVIGLQILDTGLVGTWTHSSVDYRNALSKITIMIPAKGNLDT 576
Db 480 YSHQLNVAECFL-----MDRRGTIPFFTWTHRSVDFNTIDAEKITQLPVVKAYALSS 533
Qy 577 NSKVIIEGPGHTGNLVYLO-----SQGRLEITCTPNSTQSYFIRLRYATNGAGNTLPNIS 632
Db 534 GASIEGPGTGNLLFLKSSNSIAKFKVTLNSAALLQRYRIRYAST-----TNLR 587
Qy 633 LTIIPGIVIGPPORLNTFSGTNNYNNQY--GDFGY--FQFPSTVTLPLNRPINRAD 688
Db 588 LFW-----QNSNDFLVIYINKTNWIDGDLTYQTFDFATS-----NSNMFGSGDTND 634
Qy 689 -----VSNSTIILDKIEFPI 704
Db 635 FIIGAESFVSNEKIYIDKIEFIV 658

RESULT 10

US-08-993-170A-112
; Sequence 112, Application US/08993170A
; Patent No. 6063597
; GENERAL INFORMATION:
; APPLICANT: English, Leigh H.
; APPLICANT: Brussock, Susan M.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Bryson, James W.
; APPLICANT: Kulesza, Caroline A.
; APPLICANT: Walters, Frederick S.
; APPLICANT: Slatin, Stephen L.
; APPLICANT: Von Tersch, Michael A.
; TITLE OF INVENTION: POLYPEPTIDE COMPOSITIONS TOXIC TO
; COLEOPTERAN INSECTS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,170A
FILING DATE: 18-DEC-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 112:
SEQUENCE CHARACTERISTICS:
LENGTH: 659 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-993-170A-112

Query Match 17.58; Score 676; DB 2; Length 659;

Best Local Similarity 28.0%; Pred. No. 5,1e-54;

Matches 208; Conservative 116; Mismatches 286; Indels 134; Gaps 28;

Qy 1 MNQNDNNEYEIDSHSTSPYFNNRNSNDSRYPTNNPQPLQNTNYKEWLNMCQNTQYG 60
Db 9 MNPN--RSEYDTTKVTNSLP---TNHQVPLADNPNSTLEELNYKEFLRMTADNST-- 62
Qy 61 DNFTFASADTIAAVSAGTIVSGTLLAGIGLTSISGPIGIIAISFGTLITVFWPAG 120
Db 63 ---EVLDSSTVDAVGTGIVGQILGWG-----VFPAGALTSFYQSFNLAIWPS- 110
Qy 121 EODKTVWTFIKMGEIFVDPPLTESIKQLKLTLEGPRLQLQSYNTALDDWRKLRLOAP 180
Db 111 --DADPKAFMAQVEVLIDDKIEYAKSKALAELOGLQNNFEDVYVALDSWK----- 161
Qy 181 GLPPSSALQQAULTKIRFENVHNDIFREIPGQLETYKTLTLLPIYAQAANFHLNLQOG 240
Db 162 -APVNLRSRQRIRELFSAQSHFRNSMPSPAVSKFEVLFLPTYAQAANTHLLLKDA 220
Qy 241 AELADEWNAIHPSQIEPNAGTSD--YYKLKENIPKYSNYCANTYRTGLKNLRDEPN 297
Db 221 QVFEGEWY-----SSEDIAEFYQRLKLTQYTDHCNVNWNVGLNSLRGSTY 268
Qy 298 MKWSIFNDYRRYMTITVLDTISQSLYDIKRYRDSIGGIEVKIKNELTRIETVTEINF 357
Db 269 DAWKFNRRFRREMTLVLDLIVLFPFYDVLYS-----KGVKELTRDIFTDPI-FT 319
Qy 358 RLPQLRVQPNLATMEYNLTRASPKLFSLEQFIYFTENTNFGNRL-VGISNRDAPTY--- 413
Db 320 LNALQEVGPTFSSIEHSIRKP--HLFDYLRGIEFHT-----RLRFGYSGKDSFNWWSG 370
Qy 414 -----SNTITETLYGERTSGPTTKTIRPFES---YKSVITVDRQSPVPSPIQPH 459
Db 371 NYVETRPSIGSNDTITSPFYGDKSIEPIQKL--SFDQKQVYRTIANTDIAAFPDKGI--Y 426
Qy 460 FIINQIEL-YLNGSSNNLTKYSAGGSLSNQTNTFFQPKKDCNCLVIDPGCS--PNFNN 516
Db 427 FGVTKVDFSQYDQKNET-----STQYDSKRYNGYLGQDSIDQLPETTDEPLEKA 479
Qy 517 YSHILSHFSLPTYVYVIGLQILDTGLVGTWTHSSVDYRNALSKITIMIPAKGNLDT 576
Db 480 YSHQLNVAECFL-----MDRRGTIPFFTWTHRSVDFNTIDAEKITQLPVVKAYALSS 533
Qy 577 NSKVIIEGPGHTGNLVYLO-----SQGRLEITCTPNSTQSYFIRLRYATNGAGNTLPNIS 632
Db 534 GASIEGPGTGNLLFLKSSNSIAKFKVTLNSAALLQRYRIRYAST-----TNLR 587
Qy 633 LTIIPGIVIGPPORLNTFSGTNNYNNQY--GDFGY--FQFPSTVTLPLNRPINRAD 688
Db 588 LFW-----QNSNDFLVIYINKTNWIDGDLTYQTFDFATS-----NSNMFGSGDTND 634

```
QY 689 -----VNSILIIDKIEPI 704
Db 635 FIIGAEFSVNEKIYIDKIEPIV 658

RESULT 11
US-08-993-775B-112
; Sequence 112, Application US/08993775B
; Patent No. 6077824
; GENERAL INFORMATION:
; APPLICANT: English, Leigh H.
; APPLICANT: Brussock, Susan M.
; APPLICANT: Malvar, Thomas W.
; APPLICANT: Bryson, James W.
; APPLICANT: Kulesza, Caroline A.
; APPLICANT: Walters, Frederick S.
; APPLICANT: Slatin, Stephen L.
; APPLICANT: Von Tersch, Michael A.
; TITLE OF INVENTION: METHODS FOR IMPROVING THE ACTIVITY OF
; TITLE OF INVENTION: DELTA-ENDOTOXINS AGAINST INSECT PESTS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,775B
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: MECO:150
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-993-775B-112

Query Match 17.5%; Score 676; DB 2; Length 659;
Best Local Similarity 28.0%; Pred. No. 5.1e-54;
Matches 208; Conservative 116; Mismatches 286; Indels 134; Gaps 26;

QY 1 MNQNNNDNEVEIIDSHTSPYFPNNSNDSPYTNPNQPLQNTNYKEWLNMCQNTQYG 60
Db 9 MNFNN-RSEYDTIKVTNSELPL---THNQYPLADNPENSTLEELNYKEFLRWMTADNST-- 62

QY 61 DNFETTFASADTIAAASAGTIVSGTLLAGIGLTSISGPIGIIAGIIISFGTLITVFWPAG 120
Db 63 ---EVLDSSTVKDAGVGISVVGQILGVG-----VPPAGALTSFYQSFLNAIWPS- 110

QY 121 EQDKTWTQIKGEIFVDPTPLTESIKQLKLTUEGFRQILQSNVTALDWRKUKRLQAP 180
Db 111 --DADPWKPAQVAVLIDKIEYAKSKALAEQLQNNFVYVNALDSWK----- 161

QY 181 GLPSSALQAALTKIRFENVHNDFTREIPGQLEYTKLLLPYQAANFHLNLAQQG 240
Db 162 -APVNLRSRQRIRELFQSAESHFRNMPSPFAVSKFEVFLFLPYQAANTHLLLKDA 220
```

```
QY 241 AELADEWNADIHPSQIEPNAGTSDP---YYKLLKENIKPKSYNYCANTYRTGLKNLRDEPN 297
Db 221 QVFGENG---SSEDIAEFYQRLKLTQOYTDHCVMWYVGLNSLRSSTY 268

QY 298 MKWSIFNDYRYMTITVLDTISQFSLYDIKRVDSIGGIEVGKIKNELTREIYTTTEINF 357
Db 269 DAWVKFRFRREMTLVLDLVLFFFYDRLYS-----KGVKTELTRDIFTDFI-FT 319

QY 358 RLPQLRVQPNLATMEYNLTRASFKLPSFLEQFIYFNTNTNFQNL-VGISNRDAPTY--- 413
Db 320 LNALQEGYPTFSSSIENSIRKP--HLFDYLRGIEPHT-----RLRPGYSGKDSFNWSG 370

QY 414 -----SNTITELYGERTGSPTKTIRPES---YKVSIVTDQSPSPVSIQPH 459
Db 371 NYVETRPISGNDTITSPFYGDKSIEPIQKL--SFDGQKVYRTIANTIAAFPDKGI--Y 426

QY 460 FIINQIEL-YLNGSSNNTLYKSAGGSLSNYONTTFQFPRKKCNLVIDPGCS--PNFNN 516
Db 427 FGVTKVDFPSYDDQKNET-----STQYDSKRYNGYLGADSIDQLPETTDELEKA 479

QY 517 YSHILSHPSLFYSYVIGLQILDTGVGLWTHSSVDRYNAISDKIITMIPAIGNNLDT 576
Db 480 YSHQLNVAECFL-----MQDRRGTIPTFTWTHRSVDFFENTIDAEKITQLPVVKAVALSS 533

QY 577 NSKVIIEGPGHTGMLVYLQ-----SQRLIEITCETENSTQSYFIRLRYATNGAGNTLPNIS 632
Db 534 GASIEGPGFTGGNLLFLKSSNSIAKFKVLNSAALLQRYVRIRYAST-----TNLR 587

QY 633 LTIPGVIGIPQRLNNTSPGTYNNLIQY--GDFGY--FOFPGSTVTLPLNRNIPFIFNRAD 688
Db 588 LFV-----QNSNDFLVYINKTMIDGLTYTQTFDFATS-----NSNMGFSGDTND 634

QY 689 -----VNSILIIDKIEPI 704
Db 635 FIIGAEFSVNEKIYIDKIEPIV 658

RESULT 12
US-09-427-770-112
; Sequence 112, Application US/09427770
; Patent No. 6620988
; GENERAL INFORMATION:
; APPLICANT: English, Leigh H.
; APPLICANT: Brussock, Susan M.
; APPLICANT: Malvar, Thomas W.
; APPLICANT: Bryson, James W.
; APPLICANT: Kulesza, Caroline A.
; APPLICANT: Walters, Frederick S.
; APPLICANT: Slatin, Stephen L.
; APPLICANT: Romano, Michael A.
; TITLE OF INVENTION: NUCLEIC ACID SEGMENTS ENCODING MODIFIED
; TITLE OF INVENTION: COLEOPTERAN-TOXIC CRYSTAL PROTEINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/427,770
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/993,722
; FILING DATE: 18-DEC-1997
```

```
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Kitchell, Barbara S.
/ REGISTRATION NUMBER: 33,928
/ REFERENCE/DOCKET NUMBER: MECO:149
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 512/418-3106
/ TELEFAX: 512/474-7577
/ INFORMATION FOR SEQ ID NO: 112:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 659 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/
US-09-427-770-112

Query Match 17.5%; Score 676; DB 2; Length 659;
Best Local Similarity 28.0%; Pred. No. 5.1e-54;
Matches 208; Conservative 116; Mismatches 286; Indels 134; Gaps 28;

QY 1 MNQNDNNEVEIIDSHTSPYFPNRNSNDSRYPTNNPQPLONTNYKEWLNMCQNTQYG 60
DB 9 MNPNN-RSEYDTTKVTPNSLTP---TNNQYPLADNPNSTLELNKEFLMTADNST-- 62

QY 61 DNPETASADTIAAVSAGTIVSGTLLAGIGLTSISGPIGIIAISFGTLITVFPAG 120
DB 63 ---EVLDSSTVKDVGTSVVGQILGVG-----VFPAGALTSFYQSFLNAIWPS- 110

QY 121 EODKTVMTQFIKMEIFVDPPLTESIKQLKLTLEGFQILQSVNTALDDWRKLRLOAP 240
DB 162 -APVNLRSRSDRIREFLSQAESHFNSMPSFAVSKFEVLFLPTVQAANTHLLLKDA 220

QY 241 AELADEWNADIHPSQIEPNAGTSD---YKLLKENIPKYSNYCANTYRTGLKNLRDEPN 297
DB 221 QVFGEWGY-----SSEDIAEFYQRLKLTQVYTDHCVNWNVGLNSLRGSTY 268

US-09-427-769-112

Query Match 17.5%; Score 676; DB 2; Length 659;
Best Local Similarity 28.0%; Pred. No. 5.1e-54;
Matches 208; Conservative 116; Mismatches 286; Indels 134; Gaps 28;

QY 1 MNQNDNNEVEIIDSHTSPYFPNRNSNDSRYPTNNPQPLONTNYKEWLNMCQNTQYG 60
DB 9 MNPNN-RSEYDTTKVTPNSLTP---TNNQYPLADNPNSTLELNKEFLMTADNST-- 62

QY 61 DNPETASADTIAAVSAGTIVSGTLLAGIGLTSISGPIGIIAISFGTLITVFPAG 120
DB 63 ---EVLDSSTVKDVGTSVVGQILGVG-----VFPAGALTSFYQSFLNAIWPS- 110

QY 121 EODKTVMTQFIKMEIFVDPPLTESIKQLKLTLEGFQILQSVNTALDDWRKLRLOAP 240
DB 162 -APVNLRSRSDRIREFLSQAESHFNSMPSFAVSKFEVLFLPTVQAANTHLLLKDA 220

QY 241 AELADEWNADIHPSQIEPNAGTSD---YKLLKENIPKYSNYCANTYRTGLKNLRDEPN 297
DB 221 QVFGEWGY-----SSEDIAEFYQRLKLTQVYTDHCVNWNVGLNSLRGSTY 268

US-09-427-769-112

Query Match 17.5%; Score 676; DB 2; Length 659;
Best Local Similarity 28.0%; Pred. No. 5.1e-54;
Matches 208; Conservative 116; Mismatches 286; Indels 134; Gaps 28;

QY 1 MNQNDNNEVEIIDSHTSPYFPNRNSNDSRYPTNNPQPLONTNYKEWLNMCQNTQYG 60
DB 9 MNPNN-RSEYDTTKVTPNSLTP---TNNQYPLADNPNSTLELNKEFLMTADNST-- 62

QY 61 DNPETASADTIAAVSAGTIVSGTLLAGIGLTSISGPIGIIAISFGTLITVFPAG 120
DB 63 ---EVLDSSTVKDVGTSVVGQILGVG-----VFPAGALTSFYQSFLNAIWPS- 110

QY 121 EODKTVMTQFIKMEIFVDPPLTESIKQLKLTLEGFQILQSVNTALDDWRKLRLOAP 240
DB 162 -APVNLRSRSDRIREFLSQAESHFNSMPSFAVSKFEVLFLPTVQAANTHLLLKDA 220

QY 241 AELADEWNADIHPSQIEPNAGTSD---YKLLKENIPKYSNYCANTYRTGLKNLRDEPN 297
DB 221 QVFGEWGY-----SSEDIAEFYQRLKLTQVYTDHCVNWNVGLNSLRGSTY 268
```



```

; INFORMATION FOR SEQ ID NO: 38:
;
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 652 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-996-441B-38

Query Match
Best Local Similarity 28.1%; Score 663; DB 2; Length 652;
Matches 208; Conservative 114; Mismatches 292; Indels 126; Gaps 27;

Qy 1 MNQNNDNNEYIIDSHTSPYFPNR--NSNDSRYPTYNPNQPLQNTNYKEWLNMCQ-GNT 57
Db |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy 1 MNPNN-RSEHD----TIKVTNSELSQTNHQVPLADNPNSLTLEELNYKEFLRMTEDSST 54
Db |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy 58 QYGNPFETPASADTIA-AVSAAGTVISGTLLAGLGLTSISGPTIGIICAIISFGTLITVF 116
Db |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy 55 EVLDN-----STVKDAVGTGIVSVQILGVVG-----VPPAGALTSFVQSFLNTI 99
Db |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy 117 WPAGEQDKTVMQFIKKGEIFVDTPITESTIKQLQTLQLEGFQRIQLOSINTALDDRWKLKR 176
Db |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy 100 WPS---DADPKAFMAQAVEVLIDKKIEEYAKSALAEQLQNNFEDYVNALNSWKK--- 153
Db |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy 177 LQAPGLPPSALCOQAALTILKIRENVHNDPIREIPGFOLEYTKTLTLLPIYQAANPHELNL 236
Db |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy 154 -----TPLSUKRSQDRIRLEFSQAESHFRNSMPSFAVSKFEVLFTPYQAANTHLLL 208
Db |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy 237 LQGAELADEWNADIHPISQIEPNAGTSDDYIKLLKENIPKYSNYCANTYRTGLKNLRDEP 296
Db |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy 209 LKDAQVFGESWGYG-----SEDAVEAFYHRQKLT-QQYTDHCVNWVNVGLNGLRGST 259
Db |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

```

```

260 YDAWVKFNPRREMTLTVLDLVLVFPFDRLVY-----KGVKTELTRDIFTDPIFS 311
QY 357 DRLPQLRVOPNLATMEYNLTRASFCLFSPLEQEIFYT-----ENTNF--GNRLVGI 405
Db 312 LRTP-LAYGPTFLSIENIRKP--HLFDYLOQIEFHRLQPGYFGKDSFNWNGY---V 365
QY 406 SNRDAPYTSNITETLYGBERTGSPPTKTIRPPFS---YKVISVTRDQSP--VSP 455
Db 366 ETRPSIGSSKTIITSPFGDKSTEPVQKL--SPDGQKVYRTIANTDVAAMPNGKVILGVTK 423
QY 456 IQPHFIINQIELYLVGSSNNLTLYSAGGSLSYNQNTIFFQPRKDCNCLVIDGCCSPNFN 515
Db 424 VD-----FSQYDDQKNSTQTQYDKKNNGHVSAQDSIDQLPPTTD-----BPEK 470
QY 516 NYSHILSHFSLFYTSYVIGLQILDTGVLGTHSSVDRYNAISDKIITMIPAKGNLND 575
Db 471 AYSHQLNYAECL-----WQDRGTPPTTWHRSVDFTDAEKITQLPVVKAVALS 524
QY 576 TNSKVIEGPGHTGNNLVYIQ-----SQGRLEITCETPNSTQSYPIRURYATNGAGTLPN 631
Db 525 SGASIIIEGPGCTGNNLLFLKSSNSIAKPKVTLNSAALLQRYVRIRYAST-----TNL 578
QY 632 SLTIPGVGIPPORLANTSGTNYN-----NLOYGDEGYFQFPSTVTLPLNRNIPFIF 684
Db 579 RLFP-----QNSNNDLVIYINKMTMKDDDLTYQTFDLATNNGMFGSGDKNELIIG 630
QY 685 NRADVSNLSILIIDKIEFIPI 704
Db 631 AESFVSNKXIYIDKIEFIPV 650

RESULT 15
US-08-993-722A-38
; Sequence 38, Application US/08993722A
; Patent No. 6060594
; GENERAL INFORMATION:
; APPLICANT: English, Leigh H.
; APPLICANT: Brussock, Susan M.
; APPLICANT: Malvar, Thomas M.

```

APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Tersch, Michael A.
APPLICANT: Romano, Charles
TITLE OF INVENTION: NUCLEIC ACID SEGMENTS ENCODING MODIFIED
TOXIC CRYSTAL PROTEINS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESS: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,722A
FILING DATE: 18-DEC-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:149
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3106
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 652 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-993-722A-38

Query Match 17.1%; Score 663; DB 2; Length 652;
Best Local Similarity 28.1%; Pred. No. 8.3e-53;
Matches 208; Conservative 114; Mismatches 292; Indels 126; Gaps 27;
QY 1 MNQNDNNEYIIDSHTSPYFNNR--NSNDSRYPTNNPNQIQTNTYKELNMQ--GNT 57
DB 1 MNPNN-RSEHD-----TIKVTNSELQTNHQYPLADNPNTLEELNYKEFLMTDSST 54
QY 58 QYGDNPETASADTIA-AVSAGTIVSGTLLAGIGGTSISGPIGIGAIISFGTLITVF 116
DB 55 EVLDN-----STVRDAVGTGISVVGQILGVVG-----VPFAGALTSFYQSFLNTI 99
QY 117 WPAEQDKTWTQFIKMGIFVDTPITESTIKQLQTLLEGFRQLQSYNTALDDWRKLR 176
DB 100 WFS---DADPWKAFMAQVEVLIDKKIEYAKSALAEQLQNNFEDYVNALSWKK--- 153
QY 177 LQAPGLPPSSALQQAALTILKIRENVNDPIREIPGQLETYKTLTLLPIYAQAANFHLNL 236
DB 154 -----TPLSLRGRSQDRLEFSAESHFRNMPFVSKFEVLFPTTAAQANTHLL 208
QY 237 LQQAELADEWNADIHPSQIEPNAGTSDDYKLLKENIPKSYNCANTYRTGLKLRDEP 296
DB 209 LKDAQVFGGEWGS-----SEDAEFVHRQLKT-QQYTDHCYVNVVGLNGLRGST 259
QY 297 NMKWSIFNDYRRYMTITVLDITQPSLYDIKRYRDSIGGIEVGIKNELTREIYTTINF 356
DB 260 YDAWKFNFRREMTLTVLDLIVLFPFYDIRLS-----KGVKTELTRDIFTDPIFS 311
QY 357 DRLPOLRVQPNLATMEYNLTFRASFKLPFLFOFIYV-----ENTNF--GNRLVGI 405
DB 312 LKTP-LAYGPTLUSIENSIRKP--HLFDYQLQIEFTRLPQGVFGKDSFNWSGNY---V 365
QY 406 SNRDAPYTSNTITETLYGERTGSPFTTKTRPPES---YKVSIVTDRQSP-----VSP 455

Db 366 ETRPSIGSSKTIITSPPFYGDKSTEPVQKL--SFDQKVYRTTANTDVAAMPNGKVLGVTK 423
QY 456 IQPHFIINQIELYLNSSNNTLYKSAGSLSNYQNTTFFQFPRKDCNLVIDPGGSPNPN 515
Db 424 VD---FSQYDDQKNETSTQTYDSKRNGHVSAQDSIDQLPETTD-----EPLEK 470
QY 516 NYSHILSHFSLFTYVYVIGLQQLDGTGLGTHSSVDRYNAISDKIITMIPAIKGNLND 575
Db 471 AYSHQLNYAECFL-----MODRRGTIPFFTWTHRSVDFNTIDAEEKITQLPVVKAYALS 524
QY 576 TNSKVIIEGPGHTGGLVYLQ-----SQGRLEITCETPNSTQSYFIRLRYATNGAGNTLPNI 631
Db 525 SGASIIIEGPGTGNLLFLKESNSIAKFKVTLNSAALLQRYVRIRYAST-----TNL 578
QY 632 SLATIPGVIGIPPPQRLNNTFSGTNYN-----NLQYGDGFGYFQFPSTVTLPLNRNIPFI 684
Db 579 RLFV-----QNSNNDFLVIYINKTMKDDDLTYQTFDLATTNSNMFGSGDKNELIIG 630
QY 685 NRADVNSIILIDKIEFPI 704
Db 631 AESFVSNEKIVIDKIEFIPV 650

Search completed: December 4, 2005, 12:51:47
Job time : 52 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2005, 12:46:17 ; Search time 166 Seconds
(without alignments)
1890.027 Million cell updates/sec

Title: US-10-783-417-2
Perfect score: 3869
Sequence: 1 MNQNNNVEYIDSHSPY.....KLETIQKINRFTNHTKIL 735

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.Main:*
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pap:*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pap:*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pap:*
4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pap:*
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pap:*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3869	100.0	735	5	US-10-783-417-2
2	3318	85.8	744	4	US-10-782-570-2
3	3104.5	80.2	694	4	US-10-782-570-4
4	1060.5	27.4	1180	4	US-10-782-141-12
5	1060.5	27.4	1180	4	US-10-782-096-14
6	1060.5	27.4	1180	4	US-10-782-570-10
7	1060.5	27.4	1180	5	US-10-783-417-8
8	1060.5	27.4	1180	5	US-10-781-979-14
9	892.5	23.1	675	4	US-10-782-141-17
10	892.5	23.1	675	4	US-10-782-096-18
11	892.5	23.1	675	4	US-10-782-570-14
12	892.5	23.1	675	5	US-10-783-417-12
13	892.5	23.1	675	5	US-10-781-979-19
14	892.5	23.1	675	5	US-10-926-819-16
15	849	21.9	1109	3	US-09-756-526A-4
16	849	21.9	1109	4	US-10-345-020-4
17	849	21.9	1109	4	US-10-342-821-4
18	758.5	19.6	1136	4	US-10-782-141-13
19	758.5	19.6	1136	5	US-10-781-979-15
20	758.5	19.6	1136	5	US-10-929-754-1
21	709	18.3	682	4	US-10-782-096-20
22	709	18.3	682	4	US-10-782-570-16
23	709	18.3	682	5	US-10-783-417-14
24	709	18.3	682	5	US-10-781-979-21
25	703	18.2	1210	4	US-10-032-717-4
26	703	18.2	1210	4	US-10-414-637-4
27	703	18.2	1210	4	US-10-606-320-4

28	703	18.2	1210	4	US-10-746-914-4	Sequence 4, Appli
29	689.5	17.8	1386	4	US-10-120-544A-6	Sequence 6, Appli
30	689.5	17.8	1386	6	US-11-091-654-6	Sequence 6, Appli
31	678.5	17.5	667	4	US-10-032-717-8	Sequence 8, Appli
32	678.5	17.5	667	4	US-10-414-637-8	Sequence 8, Appli
33	678.5	17.5	673	4	US-10-032-717-18	Sequence 18, Appli
34	678.5	17.5	673	4	US-10-414-637-18	Sequence 18, Appli
35	678.5	17.5	673	4	US-10-606-320-14	Sequence 14, Appli
36	678.5	17.5	673	4	US-10-746-914-14	Sequence 14, Appli
37	676	17.5	659	4	US-10-614-076-112	Sequence 112, App
38	676	17.5	659	4	US-10-782-096-12	Sequence 12, Appli
39	676	17.5	659	4	US-10-782-570-9	Sequence 9, Appli
40	676	17.5	659	5	US-10-783-417-7	Sequence 7, Appli
41	675	17.4	1206	4	US-10-032-717-2	Sequence 2, Appli
42	675	17.4	1206	4	US-10-414-637-2	Sequence 2, Appli
43	675	17.4	1206	4	US-10-606-320-2	Sequence 2, Appli
44	675	17.4	1206	4	US-10-746-914-2	Sequence 2, Appli
45	672.5	17.4	648	4	US-10-782-141-18	Sequence 18, Appli

ALIGNMENTS

RESULT 1
US-10-783-417-2
; Sequence 2, Application US/10783417
; Publication No. US20040216186A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargies, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-006, A Delta-Endotoxin Gene and
; TITLE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/274146
; CURRENT APPLICATION NUMBER: US/10783,417
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: 60/448,806
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-783-417-2

Query Match	100.0%	Score 3869;	DB 5;	Length 735;
Best Local Similarity	100.0%;	Pred. No. 1.6e-304;		
Matches 735;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MNQNNNVEYIDSHSPYFNNSNDSRYPTNNPNQPLONTNYKSWLNMCQNTQVG	60	
Db	1	MNQNNNVEYIDSHSPYFNNSNDSRYPTNNPNQPLONTNYKSWLNMCQNTQVG	60	
Qy	61	DNFETPASADTIAAVSAGTIVSGTLLAGIGLTSISGPIGIIIGAIISFGTLITVFWPAG	120	
Db	61	DNFETPASADTIAAVSAGTIVSGTLLAGIGLTSISGPIGIIIGAIISFGTLITVFWPAG	120	
Qy	121	EQDKTVMQTQIKMGEIFVDTPLTSTSIKQLKLTLEGRFQILQSNTALDDWRKLRLQAP	180	
Db	121	EQDKTVMQTQIKMGEIFVDTPLTSTSIKQLKLTLEGRFQILQSNTALDDWRKLRLQAP	180	
Qy	181	GLPSSALQQAULTKIRFENVHNDFIREIPGFOLETYKTLPIYQAANPHNLQOG	240	
Db	181	GLPSSALQQAULTKIRFENVHNDFIREIPGFOLETYKTLPIYQAANPHNLQOG	240	
Qy	241	AELEADENADHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKRLRDEPNMKW	300	
Db	241	AELEADENADHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKRLRDEPNMKW	300	
Qy	301	SIFNDYRRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKGIKNELTREIYTTTINFDRLP	360	

Db 301 SIENDYRRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKGKINELTREIYTTINFDRLP 360
QY 361 QLRVQPNLATMEYNLTRASFLEQFIFYYTENTFNGNRLVGINRDAPTYSNITITET 420
Db 361 QLRVQPNLATMEYNLTRASFLEQFIFYYTENTFNGNRLVGINRDAPTYSNITITET 420
QY 421 LYGERTGSPPTKTIIRPFESYKVSIVTDROSPPVSPIQPHPIINQIIBLYLNGSSNNLKY 480
Db 421 LYGERTGSPPTKTIIRPFESYKVSIVTDROSPPVSPIQPHPIINQIIBLYLNGSSNNLKY 480
QY 481 AGGSLSNYQNTFFQPPRKDCNVLDPGCSFNPNYSHTLSHSLFTSYVIGLOQIL 540
Db 481 AGGSLSNYQNTFFQPPRKDCNVLDPGCSFNPNYSHTLSHSLFTSYVIGLOQIL 540
QY 541 DTGVLGWTSHSSVDYRNAISDKIITMIPAKGNLDTNSKVIIEGPGHTGGNVLVLSQGR 600
Db 541 DTGVLGWTSHSSVDYRNAISDKIITMIPAKGNLDTNSKVIIEGPGHTGGNVLVLSQGR 600
QY 601 EITCETPNSTQSYFIRLRYATNGAGNTLPNISLTIPGVIGIPQRLNNTFSGTNNYNNLO 660
Db 601 EITCETPNSTQSYFIRLRYATNGAGNTLPNISLTIPGVIGIPQRLNNTFSGTNNYNNLO 660
QY 661 GDFGYFQFPSTVTLPLNRNIPFIRNADVSNLSIITDKIEFIPITSSMHQNRKQKLETI 720
Db 661 GDFGYFQFPSTVTLPLNRNIPFIRNADVSNLSIITDKIEFIPITSSMHQNRKQKLETI 720
QY 721 QTKINTFFTNHTK 735
Db 721 QTKINTFFTNHTK 735

RESULT 2

US-10-782-570-2
; Sequence 2, Application US/10782570
; Publication No. US20040210965A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-007, A Delta-Endotoxin Gene and
; FILE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/274144
; CURRENT APPLICATION NUMBER: US/10782,570
; CURRENT FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: 60/448,812
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 744
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-782-570-2

Query Match 85.8%; Score 3318; DB 4; Length 744;
Best Local Similarity 87.1%; Pred. No. 8.3e-260;
Matches 639; Conservative 31; Mismatches 60; Indels 4; Gaps 3;

QY 1 MNQNDNNEYIIDSHTSPVFPNPNDSKRYPTNNPNQPLQNTYKWLNMCCQNTQYG 60
Db 1 MNQNN-NNEYIIDSKNLSPNSRNDHRYPTNNPNQPLQNTYKWLNMCCQNTQYG 59
QY 61 DNFETPASADTTAAVSAGTIVSGTLGAGLTSISGPIGIIIGAIISFTLITVFPAG 120
Db 60 DNFETPASADTTAAVSAGTIVSGTLGAGLTSISGPIGIIIGAIISFTLITVFPAG 119
QY 121 EQDKTWTQFIKMGEIFVDTPLTESIKQLKLTLEGFRQILQSYNTALDWRKLRLOAP 180
Db 120 EQDKTWTQFIKMGEIFVDTPLTESIKQLKLTLEGFRQILQSYNTALDWRKLRLOAP 179

QY 181 GLPSSALQQAALTLLKIRFENVHNDPIREIPGFOLEYTKTLLLPYAAQANFHLNLLQOG 240
Db 180 GLPSSALQQAALTLLKIRFENVHNDPIREIPGFOLEYTKTLLLPYAAQANFHLNLLQOG 239
QY 241 AELADENWADIHPSQIEPNAGTSDYYKLLKENIPKYSNYCANTYRTGLKRLDEPNMKW 300
Db 240 AELADENWADIHPSQIEPNAGTSDYYKLLKENIPKYSNYCANTYREGINKRLNEPNRW 299
QY 301 SIENDYRRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKGKINELTREIYTTINFDRLP 360
Db 300 SIENDYRRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKGKINELTREIYTTINFDRLP 357
QY 361 QLRVQPNLATMEYNLTRASFLEQFIFYYTENTFNGNRLVGINRDAPTYSNITITET 420
Db 358 YLBQIPNLATMEYNLTRSGIRLFSFLDELIFYYTKNETYGNRLVGIANRNRSTYATTGTEI 417
QY 421 LYGERTGSPPTKTIIRPFESYKVSIVTDROSPPVSPIQPHPIINQIIBLYLNGSSNNLKY 479
Db 418 IYGERTGPPPTKTIIRPFESYKVSIVTDROSPPVSPIQPHPIINQIIBLYLNGSSNNLKY 477
QY 480 SAGGSLSNYQNTFFQPPRKDCNVLDPGCSFNPNYSHTLSHSLFTSYVIGLOQIL 539
Db 478 SAGGSLSNYQNTFFQPPRKDCNVLDPGCSFNPNYSHTLSHSLFTSYVIGLOQIL 537
QY 540 LDTGVLGWTSHSSVDYRNAISDKIITMIPAKGNLDTNSKVIIEGPGHTGGNVLVLSQGR 599
Db 538 LYTALGWTSHSSVDYRNAISDKIITMIPAKGNLDTNSKVIIEGPGHTGGNVLVLSQGR 597
QY 600 LEITCETPNSTQSYFIRLRYATNGAGNTLPNISLTIPGVIGIPQRLNNTFSGTNNYNNLO 659
Db 598 LEITCETPNSTQSYFIRLRYATNGAGNTLPNISLTIPGVIGIPQRLNNTFSGTNNYNNLO 657
QY 660 YGDFGYFQFPSTVTLPLNRNIPFIRNADVSNLSIITDKIEFIPITSSMHQNRKQKLET 719
Db 658 YGDFGYFQFPSTVTLPLNRNIPFIRNADVSNLSIITDKIEFIPITSSMHQNRKQKLET 717
QY 720 IOTKINTFFTNHTK 733
Db 718 IOTKINTFFTNHTK 731

RESULT 3

US-10-782-570-4
; Sequence 4, Application US/10782570
; Publication No. US20040210965A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-007, A Delta-Endotoxin Gene and
; FILE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/274144
; CURRENT APPLICATION NUMBER: US/10782,570
; CURRENT FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: 60/448,812
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 694
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-782-570-4

Query Match 80.2%; Score 3104.5; DB 4; Length 694;
Best Local Similarity 87.6%; Pred. No. 1.6e-242;
Matches 598; Conservative 30; Mismatches 52; Indels 3; Gaps 2;

QY 52 MCCGNTQYGNFETPASADTTAAVSAGTIVSGTLGAGLTSISGPIGIIIGAIISFT 111
Db 1 MCCGNTQYGNFETPASADTTAAVSAGTIVSGTLGAGLTSISGPIGIIIGAIISFT 60

QY 112 LITVFWPAGBQDKTWTQFIMKGEIFVDTPLTBSIKQLQTLGFRQILOSNTALDDW 171
DB 61 LITVFWPAGBQDKTWTQFIMKGEIFVDTPLTBSIKQLQTLGFRQILOSNTALDDW 120
QY 172 RKLKRLQAPGLPSSALQQAALTKIRFENVHNDFIREIFGQLETKYKTLPLPIYAQAAN 231
DB 121 RKLKRLQAPGLPSSALQQAALTKIRFENVHNDFIREIFGQLETKYKTLPLPIYAQAAN 180
QY 232 FHLNLQQAELADEWNADIHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKN 291
DB 181 FHLNLQQAELADEWNADIHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKN 240
QY 292 LRDEPNKMSIFNDYRYMTITVLDITISQFSLVDIKYRDSIGIEVKGKINELTREIYT 351
DB 241 LRDEPNKMSIFNDYRYMTITVLDITISQFSLVDIKYRDSIG--RIGGKINELTREIYT 298
QY 352 TEINFDRLPQLRVQPNLATMEYNLTRASFKLFSLEQFIFETENTNFGNRLVGINRDAP 411
DB 299 TEINFDRLTYLEIQPNLAIMEYNLTRSGLRFLSFLDELIFYTKNETYGNRLVGINRNS 358
QY 412 TVSNNTITELYGRTGSPPTKTTRPESYKVSIVTDQSPVPSPI--OPHFIIINOIYLYN 470
DB 359 TVATTGTEIYIGRTGSPPTKTTRPESYKVSIVTDQVTPSPFPNIYFTINQIYLYN 418
QY 471 GSSNNTLKYSAGSLSNQNTTFFQPPKXDCNLDVIDPGSPNPNYSHILSHFSLEFYS 530
DB 419 NSPSNKLTYSAGSLSNQNTTFFQPPKXDCNLDVIDPGSPNPNYSHILSHFSLEFYS 478
QY 531 YVIGLQILQDVLGWTSSVDYRNAISDKITIMPAIKGNLDNLSKVIKGGPHTGGN 590
DB 479 YKIGLALNLYTGALGWTSSVDYRNAISDKITIMPAIKGNLDNLSKVIKGGPHTGGN 538
QY 591 LVYLSQGRLEITCETPNSTQSYFIRLYATNAGNTLPNISLTIPGVIQPPQRLNNTF 650
DB 539 LVYLSQGRLEITCETPNSTQSYFIRLYATNAGNTLPNISLTIPGVIQPPQRLNNTF 598
QY 651 SGTNNYLNQDGYGQFSTVTPLNLRNIPFIFNRADVNSILIIDKIEFIPITSSMHQ 710
DB 599 SGTNNYLNQDGYGQFSTVTPLNLRNIPFIFNRADVNSILIIDKIEFIPITSSVRQ 658
QY 711 NREKOKLETIOTKINTFFTNHTK 733
DB 659 NREKOKLETIOTKINTFFTNHTK 681

RESULT 4
US-10-782-141-12
; Sequence 12, Application US/10782141
; Publication No. US20040197917A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: A Delta-Endotoxin Gene and
; FILE REFERENCE: 045600/274143
; CURRENT APPLICATION NUMBER: US/10782,141
; PRIOR FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: 60/448,632
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1180
; TYPE: PRN
; ORGANISM: Bacillus thuringiensis
US-10-782-141-12
Query Match 27.4%; Score 1060.5; DB 4; Length 1180;
Best Local Similarity 35.1%; Pred. No. 2.2e-76;

Matches 271; Conservative 118; Mismatches 278; Indels 105; Gaps 24;
QY 1 MNQNDNNEIHIIDSHSTPYFPNRSNDSPYTNPNQPLQNTYKWEKLNCCQNTQVG 60
DB 1 MNPNQNKNEYETLNASQKL--NISNNYTRYPIENSQKLQSTNTYKDWLNCCQNTQVG 58
QY 61 DNPFETASADTIAAASAGTIVSGTLLAGIGGTISGPGIIGAIISFGTLITVFWPAG 120
DB 59 GDFETTFDS---GELSAYTIVGVLTGFGFTT---PLGL---ALIGFTLIPVLFPQAQ 108
QY 121 EQDKTVMVTQIKMGEIFVDTPLTBSIKQLQTLGFRQILOSNTALDDMKLRLQAP 180
DB 109 DOSNT--NSDFTITQKNIIKKEIASTYISNANKILNRSFNVISTYHNHLKTWE---NNP 162
QY 181 GLPSSALQQAALTKIRFENVHNDFIREIP--GFQLEYTKTLPLPIYAQAANFHLNLQ 238
DB 163 NPQNTQDVRTQIQLVHVFQVPIPELVNSCPNPSDCDYINILVLSYAAQAANLHLTVLN 222
QY 239 QGAELADEWNADIHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKNLDEP-- 296
DB 223 QAVFEAYLKNRQFDYLEP--LPTAIDYYPVLTKAIEDYTNVCVTTYKKGLNLKITTPDS 281
QY 297 ---NMKWSIFNDYRYMTITVLDITISQFSLVDIKYRDSIGIEVKGKINELTREIYT 352
DB 282 NLDCNINWNTYRTKMTAVLDLVALFPNYDVCKY-----PIGVQSELITREIYQV 333
QY 353 EINFDRLPQLRVQPNLATMEYNLTRASFKLFSLEQFIFY-----TENTNFGNR----- 401
DB 334 -LNFEESPYYKY--DFQVQEDSLTRRP--HLFTWLDLSLNFYEKAQTTNNFFTSYNNMPHY 389
QY 402 -LVGISNRDAPTYNTITELYGRTGSPPTKTTRPESYKVSIVTDQSPVPSPIQHPH 460
DB 390 TLDNISOKSSVFGNHNVTDKL--KSLGLATNIYI-----F 422
QY 461 IINOIEL---YLNCGSSNNT-----LKYSAGGSLSNYQ--NTTFFQFPRKDC 502
DB 423 LENVISLONKYLDYNNISKMDFFITNGTRLLEKELTAGSQIYDVNKNIFGLPLKRR 482
QY 503 NLVIDPGSPNPNYSHILSHFSLEFYSYVIGLQILQDVLGWTSSVDYRNAISDKI 562
DB 483 ENQGNPTLPFYDYNYSHILSPIKSLIPATYKTQVY-----TFATHTSSVDPKNIYTHL 537
QY 563 ITMIPAIKGNLDNLSKVIKGGPHTGGNLYVLSQGRLEITCETPNSTQSYFIRLYATN 622
DB 538 TTQIPAVKANSGLTASQVWVGPGHTGGDLI--DFKDHFKITCQHSNFOQSYFIRIYASN 595
QY 623 GAGNTLPNISLTIPGVIQPPQRLNNTFSGTNNYLNQDGYGQFSTVTPLNLRNIPF 682
DB 596 GSANTRAVINLSIPGVAEL--GMAINPTFGDYTNLKYDKFOYLEFSNEVKFAPNQNL 654
QY 683 IFNRADY--SNSILIIDKIEFIPITSSMHQREKOKLETIOTKINTFFTNHTK 733
DB 655 VFNRSDYVTTVTLIDKIEFIPITRSIREDEKOKLETIVQOIINTFYANPIK 706

RESULT 5
US-10-782-096-14
; Sequence 14, Application US/10782096
; Publication No. US20040210964A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-009, A Delta-Endotoxin Gene and
; FILE REFERENCE: 045600/274148
; CURRENT APPLICATION NUMBER: US/10782,096
; PRIOR FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: 60/448,633
; NUMBER OF SEQ ID NOS: 23

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 1180
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-782-096-14

Query Match      27.4%; Score 1060.5; DB 4; Length 1180;
Best Local Similarity 35.1%; Pred. No. 2.2e-76;
Matches 271; Conservative 118; Mismatches 278; Indels 105; Gaps 24;

QY 1 MNQNDNNEYEYIIDSHTSPYFPNRSNDSRYPTNNPQPLQNTNYKWLNMCOGNTQYG 60
Db 1 MNPYQNKNEYETLNASQKLL--NISNNYTRYPIENSPKQLLOSTNYKDWLNMCOQNOQYG 58
QY 61 DNFEFASADTTAAVSAGTIVSGTLGAGIGLTSISGPIGIIICAIISFGTLITVFWPAG 120
Db 59 GDFEYFIDS--GELSAYTIWVGTLTGFGFT---PLGL---ALIGFTLIPVLFPQA 108
QY 121 EQDKTWTQFIKMGEIFVDPTLPTESIKQLQTLQLEGFRQILOSYNTALDWRKLRLOAP 180
Db 109 DOSNT--WSDFITQTKNIIKKEASTYISNANKILNRSFNVISTYHNHLKTWE-----NNP 162
QY 181 GLPPSSALQQAALTALKIRPENVHNDFIREF--GFQLETYKTLTLLPIYAAANFHLNLLQ 238
Db 163 NPQNTQDVRTQIQLVHYHFQNVIPELVNSCPPNPSCDYNNILVLSYAAQAAANLHLTVLN 222
QY 239 QCAELADEWNADIHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKNLRDEP-- 296
Db 223 QAVKEAYLKNRRQFDYLEP-LPTAIDYYPVLTKAIEDYTNVCVTTYKGLNLIKTTPDS 281
QY 297 ----NMKWSIFNDYRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKIGKNELTREIYTT 352
Db 334 -LNFEESPKYY--DFQYQEDSLTRRP-HLFTWLDLSNFEYKAKAQTTPNNFFTSYNNMFY 389
QY 402 -LVGISNRDAPYTSNTITETLYGERTGSPPTKTIKIRPESYKVSIVTDROQSPVPIQPHF 460
Db 390 TLDNISQKSSVFGNHNVTDKL--KSLGLATNIYI-----F 422
QY 461 IINQIEL---YINGSSNNT-----LKYSAGGSLSNYQ-NTTFFQFPRKDC 502
Db 423 LLNVISLDNKYLDYNNISKMDFFITNGTRLEKELTAGSGQITVDVKNIGLPIKRR 482
QY 503 NLVIDPGSPNPNYNSHLSHFSLFTSYVIGLQQLDGTGLVGTWTHSSVDYRNAISDKI 562
Db 483 ENQGNPTLPPTYDNYSHLSFKLSIPATYKTQVY-----TFAWTHSSVDPKNTIYTHL 537
QY 563 ITWIPAIGKNNLDTSKVIETGPGHTGGLNVLVLOSQGRLEITCETPNSTQSYFIRLYATN 622
Db 538 TTQIPAVKANSIGTASGVVQGGHTGGDLI--DFKDHFKITCOHSNFQOQSYFIRIYASN 595
QY 623 GAGNTLPNLSLTPGVIGIPPPQRLNNTFSGTNNYNNLOYGDFGFPQFPSTVTLPLNRNIPP 682
Db 596 GSANTRAVINLSIPGVAEL--GMALNPTFSGTDYTNLKYKDFQYLEFNEVKFAPNQISL 654
QY 683 IFNRADV--SNSIILIDKIEFIPITSSMHQNRKQKLETTQTKINTFTNHTK 733
Db 655 VFNRSVDYNTTTLVIDKIEFLPITRSIREDRKQKLETVQQTINTFYANPIK 706
```

RESULT 6

```
US-10-782-570-10
; Sequence 10, Application US/10782570
; Publication No. US20040210965A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
```

```
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-007, A Delta-Endotoxin Gene and
; FILE REFERENCE: 045600/27414
; CURRENT APPLICATION NUMBER: US/10/782,570
; CURRENT FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: 60/448,812
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1180
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-782-570-10
```

```
Query Match      27.4%; Score 1060.5; DB 4; Length 1180;
Best Local Similarity 35.1%; Pred. No. 2.2e-76;
Matches 271; Conservative 118; Mismatches 278; Indels 105; Gaps 24;

QY 1 MNQNDNNEYEYIIDSHTSPYFPNRSNDSRYPTNNPQPLQNTNYKWLNMCOGNTQYG 60
Db 1 MNPYQNKNEYETLNASQKLL--NISNNYTRYPIENSPKQLLOSTNYKDWLNMCOQNOQYG 58
QY 61 DNFEFASADTTAAVSAGTIVSGTLGAGIGLTSISGPIGIIICAIISFGTLITVFWPAG 120
Db 59 GDFEYFIDS--GELSAYTIWVGTLTGFGFT---PLGL---ALIGFTLIPVLFPQA 108
QY 121 EQDKTWTQFIKMGEIFVDPTLPTESIKQLQTLQLEGFRQILOSYNTALDWRKLRLOAP 180
Db 109 DOSNT--WSDFITQTKNIIKKEASTYISNANKILNRSFNVISTYHNHLKTWE-----NNP 162
QY 181 GLPPSSALQQAALTALKIRPENVHNDFIREF--GFQLETYKTLTLLPIYAAANFHLNLLQ 238
Db 163 NPQNTQDVRTQIQLVHYHFQNVIPELVNSCPPNPSCDYNNILVLSYAAQAAANLHLTVLN 222
QY 239 QCAELADEWNADIHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKNLRDEP-- 296
Db 223 QAVKEAYLKNRRQFDYLEP-LPTAIDYYPVLTKAIEDYTNVCVTTYKGLNLIKTTPDS 281
QY 297 ----NMKWSIFNDYRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKIGKNELTREIYTT 352
Db 282 NLDGNNMNNTYRTKMTTAVLDLVALFPNDVGKY-----PIGVQSELTREIYQV 333
QY 353 EINFRLPOLRVQPNLATWMEYNLTRASFLEQFIY-----TENTNFGNR----- 401
Db 334 -LNFEESPKYY--DFQYQEDSLTRRP-HLFTWLDLSNFEYKAKAQTTPNNFFTSYNNMFY 389
QY 402 -LVGISNRDAPYTSNTITETLYGERTGSPPTKTIKIRPESYKVSIVTDROQSPVPIQPHF 460
Db 390 TLDNISQKSSVFGNHNVTDKL--KSLGLATNIYI-----F 422
QY 461 IINQIEL---YINGSSNNT-----LKYSAGGSLSNYQ-NTTFFQFPRKDC 502
Db 423 LLNVISLDNKYLDYNNISKMDFFITNGTRLEKELTAGSGQITVDVKNIGLPIKRR 482
QY 503 NLVIDPGSPNPNYNSHLSHFSLFTSYVIGLQQLDGTGLVGTWTHSSVDYRNAISDKI 562
Db 483 ENQGNPTLPPTYDNYSHLSFKLSIPATYKTQVY-----TFAWTHSSVDPKNTIYTHL 537
QY 563 ITWIPAIGKNNLDTSKVIETGPGHTGGLNVLVLOSQGRLEITCETPNSTQSYFIRLYATN 622
Db 538 TTQIPAVKANSIGTASGVVQGGHTGGDLI--DFKDHFKITCOHSNFQOQSYFIRIYASN 595
QY 623 GAGNTLPNLSLTPGVIGIPPPQRLNNTFSGTNNYNNLOYGDFGFPQFPSTVTLPLNRNIPP 682
Db 596 GSANTRAVINLSIPGVAEL--GMALNPTFSGTDYTNLKYKDFQYLEFNEVKFAPNQISL 654
QY 683 IFNRADV--SNSIILIDKIEFIPITSSMHQNRKQKLETTQTKINTFTNHTK 733
Db 655 VFNRSVDYNTTTLVIDKIEFLPITRSIREDRKQKLETVQQTINTFYANPIK 706
```

```
RESULT 7
US-10-783-417-8
; Sequence 8, Application US/10783417
; Publication No. US20040216186A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargies, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-006, A Delta-Endotoxin Gene and
; TITLE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/274146
; CURRENT APPLICATION NUMBER: US/10783.417
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: 60/448,806
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1180
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-783-417-8

Query Match          27.4%; Score 1060.5; DB 5; Length 1180;
Best Local Similarity 35.1%; Pred. No. 2.2e-76;
Matches 271; Conservative 118; Mismatches 278; Indels 105; Gaps 24;

Qy      1 MNQNNDNNEYEIIIDSHTSPPYFNRNSNDSRYPTNNPQPLQNTNYKEWLNMCQNTQY 60
Db      1 MNPYQNKNEYETLNASOKKL--NISNNYTRYPIENSQKLLQSTNYKDWLNMCQNTQY 58

Qy      61 DNFTFASADTIAAVSAGTIVSGTLLAGIGLTSISGPIGIIIGAIISFGTLITVFWPAG 120
Db      59 GDETFIDS---GELSAYTIVGTVLTGFGFTT---PLGL---ALIGFGTLIPVLFPQA 108

Qy      121 EODKTVMTQIKMGEIFVDTPLTESI KQLKLTLEGFRQILQSNTALDWRKLELOAP 180
Db      109 DQSNLT-WSDFITQTKNIIKKEIASTYISNANKILNRSFNVIYHNHLKWE-----NNP 162

Qy      181 GLPPSSALQQAALTALKIRFENVHNDFFIREIP--GFQLETYKTLTLLPIYAAANFHLNLQ 238
Db      163 NPQNTQDVRTQIQLVHVFQNVIPELVNSCPNPSDCDYNNILVSSYAAANLHLTVLN 222

Qy      239 QGAELADENWADTHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKNLRDEP-- 296
Db      223 QAVKFEAYLKNRQFDYLEP-LPTAIDYYPVLTKAIEDYNYCVTYTKGGLNLKLTTPDS 281

Qy      297 ----NMKWSIFNDYRRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKGIKNELTREIYTT 352
Db      282 NLDGNINWNTYTRYTKMTTAVLDLVALFPNYDVGY-----PIGVQSELTREIYQV 333

Qy      353 EINFDRPLQRLVQPNLATMEYNLTRASPKLSFLEQIFY-----TENTNFGNR----- 401
Db      334 -LNFEESPKYKY--DFQYQEDSLTRP-HLFTWLDLSNLFYEKAQTPNNFTSHYMFHY 389

Qy      402 -LVGIGNRADPTYNTITETLYGERTSGPTTKIRPFESYKVSIVTDROSPPVSPIQPHF 460
Db      390 TLDNISQKSSVFGNHNVDKLT--KSLGLATNIYI-----F 422

Qy      461 IINQIEL---YLNQSSNNT-----LKYSAGGSLSNYQ--NTTFFQPPRKKDC 502
Db      423 LLNVISLDNKYLNDYNNISKMDFFITNGTRLLLEKELTAGSGQIYDYNKIFGLPILKRR 482

Qy      503 NLVIDGCSNPNFNYSILSHFSLFYVYVIGLQILDGVLGWTHSSVDYRNAISDKI 562
Db      483 ENQGNPLTFPTYDNYSHLSFIKSLSPATYKTVQV-----TFATWTHSSVDPKNTIYTHL 537

Qy      563 ITWIPAIGKNLDTNSKVISGPGHTGNLVYLSQGRLEITCTETPNSTQSYPIRLVATN 622
Db      538 TTQIPAVKANSLGTSKAVQVQPGHTGDLI--DPKDHFKITCQHSNFPQSYFIRIVASN 595
```

```
Qy      623 GAGNTLPNLSLTIRGVIGIPQRLNNTSGTNNYNNLOYGDFGYFQFPSPSTVTLPLNRNIPF 682
Db      596 GSANTRAVINLSIPGVAEL-GMALNPTSGDYTNLKYDFOYLFBSNEVKAPNQNISL 654

Qy      683 IFNRADV-SNSILIIDKIEFIPITSSMHQNRKQKLETIQTKINTFFTHHTK 733
Db      655 VFNRSVDVYNTTVLIDKIEFLPITRSIREDEKQKLETVOQIINTFYANPIK 706

RESULT 8
US-10-781-979-14
; Sequence 14, Application US/10781979
; Publication No. US20040250311A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargies, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-008, A Delta-Endotoxin Gene and
; TITLE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/274147
; CURRENT APPLICATION NUMBER: US/10781.979
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: 60/448,797
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 1180
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-781-979-14

Query Match          27.4%; Score 1060.5; DB 5; Length 1180;
Best Local Similarity 35.1%; Pred. No. 2.2e-76;
Matches 271; Conservative 118; Mismatches 278; Indels 105; Gaps 24;

Qy      1 MNQNNDNNEYEIIIDSHTSPPYFNRNSNDSRYPTNNPQPLQNTNYKEWLNMCQNTQY 60
Db      1 MNPYQNKNEYETLNASOKKL--NISNNYTRYPIENSQKLLQSTNYKDWLNMCQNTQY 58

Qy      61 DNFTFASADTIAAVSAGTIVSGTLLAGIGLTSISGPIGIIIGAIISFGTLITVFWPAG 120
Db      59 GDETFIDS---GELSAYTIVGTVLTGFGFTT---PLGL---ALIGFGTLIPVLFPQA 108

Qy      121 EODKTVMTQIKMGEIFVDTPLTESI KQLKLTLEGFRQILQSNTALDWRKLELOAP 180
Db      109 DQSNLT-WSDFITQTKNIIKKEIASTYISNANKILNRSFNVIYHNHLKWE-----NNP 162

Qy      181 GLPPSSALQQAALTALKIRFENVHNDFFIREIP--GFQLETYKTLTLLPIYAAANFHLNLQ 238
Db      163 NPQNTQDVRTQIQLVHVFQNVIPELVNSCPNPSDCDYNNILVSSYAAANLHLTVLN 222

Qy      239 QGAELADENWADTHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKNLRDEP-- 296
Db      223 QAVKFEAYLKNRQFDYLEP-LPTAIDYYPVLTKAIEDYNYCVTYTKGGLNLKLTTPDS 281

Qy      297 ----NMKWSIFNDYRRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKGIKNELTREIYTT 352
Db      282 NLDGNINWNTYTRYTKMTTAVLDLVALFPNYDVGY-----PIGVQSELTREIYQV 333

Qy      353 EINFDRPLQRLVQPNLATMEYNLTRASPKLSFLEQIFY-----TENTNFGNR----- 401
Db      334 -LNFEESPKYKY--DFQYQEDSLTRP-HLFTWLDLSNLFYEKAQTPNNFTSHYMFHY 389

Qy      402 -LVGIGNRADPTYNTITETLYGERTSGPTTKIRPFESYKVSIVTDROSPPVSPIQPHF 460
Db      390 TLDNISQKSSVFGNHNVDKLT--KSLGLATNIYI-----F 422

Qy      461 IINQIEL---YLNQSSNNT-----LKYSAGGSLSNYQ--NTTFFQPPRKKDC 502
Db      423 LLNVISLDNKYLNDYNNISKMDFFITNGTRLLLEKELTAGSGQIYDYNKIFGLPILKRR 482
```


Db 212 NLLKHAATYNIW---LQNGINPSTFNSNYQGLKRIQEQYDYCIQYNAAGLTWIR 268
Qy 294 DEPNMKSIENDRYRYMTITVLDTSIQSLYDIKRYRDSIGGIEVGKIKNELTRIYYTTE 353
Db 269 TTNATWNTYRYLEMTLTVLDLIAIPPNYDPEKY-----PIGVKSELIREVY-TN 319
Qy 354 INFRLPQLRVQPNLATMEYNLTIRASPKLFSLEQFIYFYTENTN-----FGNRL 402
Db 320 VNSDTF-----RTITELNGLTR-NPTLFTWINGRFRYTRNSRDILDPYDFISFTGNOM 372
Qy 403 VGISNRDAPYNTITETLYGERTG---SPTTKTIRPPESY---KVSIVTDROSPPVSP 455
Db 373 -----AFTHNDNRNIWGAVHGNIIISQDTSKVFPYRNKPIDKVEIVRREYSDI- 424
Qy 456 IQPHFIINQIELYNGSNNTLYKSAGSL-SNYQNTFFQFPKCKDCNLVIDPGCSPNF 514
Db 425 -----YEMIFPNSSEVFYSSNSTIENNYKRTDSYMPKQTKN----- 464
Qy 515 NNYSHILSHPSLFTYSYVIGLQQLDGLVGLWTHSSVDYRNAISDKIITMIPAKGNL 574
Db 465 EYGHITLSYIKTDNYIFSVVRRRV---AFSWHTSVDFTQNTIDLDNITQIHAKALKV 521
Qy 575 DTNSKVIKVGPGHTGGLVILKDSMDFRVRF-LKNVSRQYQVRIYATNA-----PKTTVF 575
Db 522 SSKSVKVGPGHTGGLVILKDSMDFRVRF-LKNVSRQYQVRIYATNA-----PKTTVF 575
Qy 635 IPGVIGIPQRLNNTFSGTNNY--NLQYDGFYGFQFPSTV-----TLPLNRNIPFI 683
Db 576 LTGIDTISVE-LPSTTSRQNPATDLTYADFGYVTFPRTVPNKTFEGEDTLMT-----L 629
Qy 684 FNRADVNSILIDIKIEFIPITSSMHQNRKQKLETTQTKINTFTTN 730
Db 630 YGTPNHSYNI-YIDKIEFIPITQSVDLYTEKQNIKTQKIVNDLFVN 675

RESULT 11

US-10-782-570-14
; Sequence 14, Application US/10782570
; Publication No. US20040210965A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Kozziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-007, A Delta-Endotoxin Gene and
; FILE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/274144
; CURRENT APPLICATION NUMBER: US/10782,570
; PRIOR FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: 60/448,812
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-782-570-14

Query Match 23.1%; Score 892.5; DB 4; Length 675;
Best Local Similarity 31.9%; Pred. No. 4e-63;
Matches 245; Conservative 126; Mismatches 267; Indels 129; Gaps 30;

Qy 1 MNQNDNNEVEIIDSHTSPYFPNRSNDSPYPTNNQPLQNTYKELWLMCCQNTQYQ 60
Db 1 MNPYQNKVEYIFNAPNGF--SKSNYSRYPLANKNPQPLKNTYKDWLNVCDNQYQ 58
Qy 61 DNFETPASADTIAAVSAGTIVSGLLAGLGLTSGISGPIGIIIGAILIISFGTLITVFPAG 120
Db 59 NNAGNFASSETIVGVSAGIIVVGTMGLG-----AFAAP--VLAAGIISFGTLIPFW-QG 109

Qy 121 EODKTVTQFIKMGEIFVDTPLTE---SIKQLKQLTLEGFRQILQSYNTALDDWRKLKRL 177
Db 110 SDPANWQDLNIG---GRPIQEDKNIIINVLTSIVTFIKNQLDKYQBFDFKWEPAF-- 163
Qy 178 QAPGLPPSSALQQAALTILKIRFENVHN---DFIREIPGFOLEYTKTLPIYAQAANFHL 234
Db 164 -----THANAKAVHDLFTTLEPIIDKDLMLKNNASVRIPT-----LPAYAQIATWHL 211
Qy 235 NLLQGAELADEWNADIHPSQIEPNAGTDDYYK--LLKENIPKYSNYCANTYRTGLKNLR 293
Db 212 NLLKHAATYNIW---LQNGINPSTFNSNYQGLKRIQEQYDYCIQYNAAGLTWIR 268
Qy 294 DEPNMKSIENDRYRYMTITVLDTSIQSLYDIKRYRDSIGGIEVGKIKNELTRIYYTTE 353
Db 269 TTNATWNTYRYLEMTLTVLDLIAIPPNYDPEKY-----PIGVKSELIREVY-TN 319
Qy 354 INFRLPQLRVQPNLATMEYNLTIRASPKLFSLEQFIYFYTENTN-----FGNRL 402
Db 320 VNSDTF-----RTITELNGLTR-NPTLFTWINGRFRYTRNSRDILDPYDFISFTGNOM 372
Qy 403 VGISNRDAPYNTITETLYGERTG---SPTTKTIRPPESY---KVSIVTDROSPPVSP 455
Db 373 -----AFTHNDNRNIWGAVHGNIIISQDTSKVFPYRNKPIDKVEIVRREYSDI- 424
Qy 456 IQPHFIINQIELYNGSNNTLYKSAGSL-SNYQNTFFQFPKCKDCNLVIDPGCSPNF 514
Db 425 -----YEMIFPNSSEVFYSSNSTIENNYKRTDSYMPKQTKN----- 464
Qy 515 NNYSHILSHPSLFTYSYVIGLQQLDGLVGLWTHSSVDYRNAISDKIITMIPAKGNL 574
Db 465 EYGHITLSYIKTDNYIFSVVRRRV---AFSWHTSVDFTQNTIDLDNITQIHAKALKV 521
Qy 575 DTNSKVIKVGPGHTGGLVILKDSMDFRVRF-LKNVSRQYQVRIYATNA-----PKTTVF 575
Db 522 SSKSVKVGPGHTGGLVILKDSMDFRVRF-LKNVSRQYQVRIYATNA-----PKTTVF 575
Qy 635 IPGVIGIPQRLNNTFSGTNNY--NLQYDGFYGFQFPSTV-----TLPLNRNIPFI 683
Db 576 LTGIDTISVE-LPSTTSRQNPATDLTYADFGYVTFPRTVPNKTFEGEDTLMT-----L 629
Qy 684 FNRADVNSILIDIKIEFIPITSSMHQNRKQKLETTQTKINTFTTN 730
Db 630 YGTPNHSYNI-YIDKIEFIPITQSVDLYTEKQNIKTQKIVNDLFVN 675

RESULT 12

US-10-783-417-12
; Sequence 12, Application US/10783417
; Publication No. US20040216186A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Kozziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-006, A Delta-Endotoxin Gene and
; FILE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/274146
; CURRENT APPLICATION NUMBER: US/10783,417
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: 60/448,806
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-783-417-12

Query Match 23.1%; Score 892.5; DB 5; Length 675;
Best Local Similarity 31.9%; Pred. No. 4e-63;
Matches 245; Conservative 126; Mismatches 267; Indels 129; Gaps 30;

```
QY 1 MNNQNNNEVEIIDSHTSPYFPNNSNDSPYPTNNPQPLONTNPKYKWLNMCGNTQYG 60
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 MNPYQNKNEYEIFNAPSNF--SKSNYSRYPLANKENQPLKNTNYKDWLNVCQDNQOYG 58
QY 61 DNFETASADTIAAVSAGTIVSGTLLAGIGLTSISGPIGIIAGIISFGTLITVFWPAG 120
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
59 NNAGNFASSETIVGVSAGIIVGTMLG-----AFAAP--VLAAGIISFGTLIPFW-QG 109
QY 121 EQDKTVMQFIKMGEIFVDTPLTE---SIKOLKLTLEGFRQILOSNTALDDWRKLR 177
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
110 SDPANWQDLNLNIG---GRPIQEIKNIIINVLTISVTPINKQDKQEFDFKWEPA-- 163
QY 178 QAPGLPSSALQQAALTLKIRFENVH---DFIREIPGFOLEYTKTLPIYAQAANFHL 234
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
164 -----THANAKAVHDLFTTLEPIIDKOLDMLKNNASYRIPT-----LPAYAQIATWHL 211
QY 235 NLLQQAELADEWADIHPSQIEPNAGTSDDYK--LKENIPKYSNCANTYRTGLKNLR 293
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
212 NLLKHAATYNIW---LQNGINPSTFNSSNYQGYLKRKIQEYDYCIQTYNAGLTMR 268
QY 294 DEPNMKWSIFNDVRRYMTIVLDTISQFSLYDIKRYRDSIGGIEVKIGIKNELTREIYTT 353
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
269 TTNATWNTYRLEMTLVLDLIAIFPNYDEKY-----PIGVKSELIREVY-TN 319
QY 354 INFDRLPQLRVQPNLATMEYNLTASPKLFSLEQFIIFYTENTN-----FGNRL 402
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
320 VNSDTF-----RTITELENGLTR-NPTLFTWINGRFFYTRNSRDILDPYDIFSFTGNQM 372
QY 403 VGISNRDAPYSNTITETLYGERTG---SPTTKTIRPFESY---KVSIVTDROSPPVSP 455
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
373 -----AFTHNDRNIIWGAVHGNIIISQTSKVPFPYRNKPIDKVEIVRHRYSDI-- 424
QY 456 IQPHFIINOIELYNGSSNNTLYKSAGGSL--SNYQNTTFFQPRKKDCNLVIDPGCSPNF 514
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
425 -----YEMIFFNSSEVFRYSNSTIENNYKRTDSYMIKQTKWN----- 464
QY 515 NNYSHILSHFSLFTYSYVIGLQIQLDGTGLVTHSSVDRYNAISDKIITWIPAIGKNL 574
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
465 EYGHGTLISYIKTDNYIYFSVVRERRV---AFSWHTSVDFQNTIDLDNIITQIHALKALKV 521
QY 635 IPGVIGIPQRLNNTFSGTNYN--NLQYGDGYFPQPSV-----TLPLNRNIPPI 683
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
576 LTGIDTISVE-LPSTTSRQPNATDLTYADFGYVTPRTVPNKTFEGEDTLLMT-----L 629
QY 684 FNRADVNSILIIDKIEFIPITSSMHQNRKQKLETIQTKINTFTFN 730
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
630 YGTPNHSYNI-YIDKIEFIPITQSVLDYTEKQIEKTKQIVNDLNVN 675
```

RESULT 13

```
US-10-781-979-19
; Sequence 19, Application US/10781979
; Publication No. US20040250311A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMT-008, A Delta-Endotoxin Gene and
; TITLE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/274147
; CURRENT APPLICATION NUMBER: US/10/781,979
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: 60/448,797
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 19
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-781-979-19
```

```
Query Match 23.1%, Score 992.5, DB 5, Length 675,
Best Local Similarity 31.9%, Pred. No. 4e-63,
Matches 245, Conservative 126, Mismatches 267, Indels 129, Gaps 30;
```

```
QY 1 MNNQNNNEVEIIDSHTSPYFPNNSNDSPYPTNNPQPLONTNPKYKWLNMCGNTQYG 60
Db 1 MNPYQNKNEYEIFNAPSNF--SKSNYSRYPLANKENQPLKNTNYKDWLNVCQDNQOYG 58
QY 61 DNFETASADTIAAVSAGTIVSGTLLAGIGLTSISGPIGIIAGIISFGTLITVFWPAG 120
Db 59 NNAGNFASSETIVGVSAGIIVGTMLG-----AFAAP--VLAAGIISFGTLIPFW-QG 109
QY 121 EQDKTVMQFIKMGEIFVDTPLTE---SIKOLKLTLEGFRQILOSNTALDDWRKLR 177
Db 110 SDPANWQDLNLNIG---GRPIQEIKNIIINVLTISVTPINKQDKQEFDFKWEPA-- 163
QY 178 QAPGLPSSALQQAALTLKIRFENVH---DFIREIPGFOLEYTKTLPIYAQAANFHL 234
Db 164 -----THANAKAVHDLFTTLEPIIDKOLDMLKNNASYRIPT-----LPAYAQIATWHL 211
QY 235 NLLQQAELADEWADIHPSQIEPNAGTSDDYK--LKENIPKYSNCANTYRTGLKNLR 293
Db 212 NLLKHAATYNIW---LQNGINPSTFNSSNYQGYLKRKIQEYDYCIQTYNAGLTMR 268
QY 294 DEPNMKWSIFNDVRRYMTIVLDTISQFSLYDIKRYRDSIGGIEVKIGIKNELTREIYTT 353
Db 269 TTNATWNTYRLEMTLVLDLIAIFPNYDEKY-----PIGVKSELIREVY-TN 319
QY 354 INFDRLPQLRVQPNLATMEYNLTASPKLFSLEQFIIFYTENTN-----FGNRL 402
Db 320 VNSDTF-----RTITELENGLTR-NPTLFTWINGRFFYTRNSRDILDPYDIFSFTGNQM 372
QY 403 VGISNRDAPYSNTITETLYGERTG---SPTTKTIRPFESY---KVSIVTDROSPPVSP 455
Db 373 -----AFTHNDRNIIWGAVHGNIIISQTSKVPFPYRNKPIDKVEIVRHRYSDI-- 424
QY 456 IQPHFIINOIELYNGSSNNTLYKSAGGSL--SNYQNTTFFQPRKKDCNLVIDPGCSPNF 514
Db 425 -----YEMIFFNSSEVFRYSNSTIENNYKRTDSYMIKQTKWN----- 464
QY 515 NNYSHILSHFSLFTYSYVIGLQIQLDGTGLVTHSSVDRYNAISDKIITWIPAIGKNL 574
Db 465 EYGHGTLISYIKTDNYIYFSVVRERRV---AFSWHTSVDFQNTIDLDNIITQIHALKALKV 521
QY 575 DTMSKVIEGPGHTGGNLVYLSQGRLEITCETPNSTQSYFIRLRVATNGAGNTLPNISLT 634
Db 522 SSOSKIVKPGHGTGGDLVILKDSMDFRVP--LKNVSRQYQVRIYATNA-----PKTTFV 575
QY 635 IPGVIGIPQRLNNTFSGTNYN--NLQYGDGYFPQPSV-----TLPLNRNIPPI 683
Db 576 LTGIDTISVE-LPSTTSRQPNATDLTYADFGYVTPRTVPNKTFEGEDTLLMT-----L 629
QY 684 FNRADVNSILIIDKIEFIPITSSMHQNRKQKLETIQTKINTFTFN 730
Db 630 YGTPNHSYNI-YIDKIEFIPITQSVLDYTEKQIEKTKQIVNDLNVN 675
```

RESULT 14

```
US-10-926-819-16
; Sequence 16, Application US/10926819
; Publication No. US20050049410A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
```

; TITLE OF INVENTION: AXMI-003, A Delta-Endotoxin Gene and
; FILE REFERENCE: Methods for Its Use
; CURRENT APPLICATION NUMBER: US/10/926,819
; CURRENT FILING DATE: 2004-08-26
; PRIOR APPLICATION NUMBER: 60/498,518
; PRIOR FILING DATE: 2003-08-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-926-819-16

Query Match 23.1%; Score 892.5; DB 5; Length 675;
Best Local Similarity 31.9%; Pred. No. 4e-63;
Matches 245; Conservative 126; Mismatches 267; Indels 129; Gaps 30;

QY 1 MNQNNNVEYIIIDSHTSPPFPNRSNDSRYPTNNNOPLONTNYKEWLNMCQNTQYG 60
Db 1 MNPYQNKNEYIEFNAPNGF--SKSNYSRYPLANKNPQKNTNYKDWLNVQCDNQY 58

QY 61 DNFETASADTAAVASGTVSGTLLAGIGLTSISGPIGIIIGAIISFGTLITVFWPAG 120
Db 59 NNAGNFASSETIVGVSGAGIIVGTMLG-----AFAAP--VLAAGIISFGTLLPIFW-QG 109

QY 121 EODKTVMTOFKMGEIFVDTPLTE--SIKQLKLTLEGFRQILOSNTALDDWRKRL 177
Db 110 SDPANWQDLLNIG---GRPIQIDKNIINVLTSIVTPIKNQDKQEPFDKWEPA-- 163

QY 178 QAPGLPSSALQQAALTKIRFENVH---DFTREIPGQLETYKTLILLPIYAAAHFL 234
Db 164 -----THAKAVHDLFTLEPIIDKOLDKNNASRIPT-----LPAYAIATWHL 211

QY 235 NLLQOGAELADENADHPQIEPNAGTSDDYK-LKKENIPKSYNCANTYRGLKNLR 293
Db 212 NLLKHAATYNIW---LONQINPSTPNSSNYQGYLKRKIQEYTDYCIQTYNAGLTMR 268

QY 294 DEPNKWSIFNDYRRYNTIIVDTISOFSLYDIKRYRDSIGGIEVKGIKNELTREIYTE 353
Db 269 TWTNATWNTYRLEMTLVLDLIAFPNYDEPKY-----PIGVKSELIREVY-TN 319

QY 354 INFDRLPQURVQNLMATMEYNLTRASPKLSFLEQFIIFYTENTN-----FGNRL 402
Db 320 VNSDTP-----RTITELNGLTR-NPTLFTWINGRFRYTRNSRDILDPYDIFSGTGNQM 372

QY 403 VGISNRDAPYSNTITETLYGERTG---SPTTKTIRPFESY---KYSIVTDROSPPVSP 455
Db 373 -----AFHTNDNRNIIWGAHVHNTISQDTSKVFPFPRNKPDKVEIYVRRHREYSII- 424

QY 456 IQPHFIINQIELYLGSSNNNTKYSAGGSL-SNYQNTTFFQFPRKDCNLVIDPGCSPNF 514
Db 425 -----YEMIFFNSSESVFYSNSTIENNYKRTDSTWIPKQTKN----- 464

QY 515 NNYSHILSHFSFTYSYVIGLQILDTVLGWTSHSSVDRYNAISDKIITMIPAIGNNL 574
Db 465 EEEYHTLSYIKTDNYIFSVVRERRV---AFSWHTTSVDQNTIDLDNIITQIHALKALKV 521

QY 575 DTNKSIVIEGHTGGNLYVYOSQGRLEITCETPNSTQSYFIRLYATNGAGNTLPNLSLT 634
Db 522 SSDSKIVKVGPGHTGGDLVILKDSMDFRVRP-LKNVSRQYQVRIYATNA-----PKTTVF 575

QY 635 IPGVIGIPPORLANTTSGTNYN--NLOYGDFGYFQFPSTV-----TLPLNRNIPFI 683
Db 576 LTGIDITISVE-LPSTTSRQPNATDLTYADFGYVTFPRTPNKTFEGEDTLMT-----L 629

QY 684 ENRADVSNSTIIDKIEFIPITSMHQREKOKLETIQTINTFTTN 730
Db 630 YGTPNHSYNI-YIDKIEFIPITOSVLDYTEKQIEKTQKIVNDLFWN 675

RESULT 15

US-09-756-526A-4
; Sequence 4, Application US/09756526A
; Patent No. US20020038005A1
; GENERAL INFORMATION:
; APPLICANT: Jana, Wojciechowska
; APPLICANT: Evgeny, Lewitin
; APPLICANT: Ludmila, Revina
; APPLICANT: Igor, Zalunin
; APPLICANT: Galina, Chestukhina
; TITLE OF INVENTION: NOVEL DELTA-ENDOTOXINS AND NUCLEIC ACID SEQUENCES CODING THEREFOR,
; FILE REFERENCE: S-30913A
; CURRENT APPLICATION NUMBER: US/09/756,526A
; CURRENT FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: US 60/175,158
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 4
; LENGTH: 1109
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-09-756-526A-4

Query Match 21.9%; Score 849; DB 3; Length 1109;
Best Local Similarity 33.3%; Pred. No. 2.9e-59;
Matches 229; Conservative 106; Mismatches 248; Indels 104; Gaps 25;

QY 68 SADTIAVASGATIVSGTLLAGIGLTSISGPIGIIIGAIISFGTLITVFWPAGQD-KTV 126
Db 28 SSDIVAVVSAGIVVVGTTLT--APASFNP-----GVVLISFGTLAPVLPDPEEDPKKI 80

QY 127 WTQFIKXGEIFVDTPLTESIKQLKLTLEGFRQILOSNTALDDWRKRLQAQGLPPSS 186
Db 81 WSPQWKHGEDLLNQTISTAVKEIALAHLNGFKDVLTYTVERAFNDWKR-----NPSA 131

QY 187 ALQQAALTKIRFENVHNDPIREIPGFOLEYKTLILLPIYAAAHFLNLLQGAELADE 246
Db 132 ---NTARLVSORFENAHFNFSVNNPQLQPTVDTLLSCYTEAANLHLNLLHQGVQFADQ 188

QY 247 WNADIHPISOIEPNAGTSDDYKLLKENIPKSYNCANTYRGLKNLRDPNNKWSIFNDY 306
Db 189 WNADQSPHMLKSSGT---YYDELLVYIEKYNICTYHKHGLHKLKSEKITWDAYNTY 245

QY 307 RRYMTITVLDTISOFSLYDIKRYRDSIGGIEVKGIKNELTREIYTEINFDRLPQURVOP 366
Db 246 RREMTLIVLDVATFPFYDIRRF-----PRGVELELTREVYTSLDHLTRPP----- 291

QY 367 NLATMEYNLTRASPKLSFLEQFIIFYTENTWFGNPLVGISNRDAPYSNT---ITETLYG 423
Db 292 -----GLFTWLSDIELYTESVAGDYLSGI--RESKYTYGNGQFTWKNIY 335

QY 424 ERTG-SPTTKTIRPFESYKYSIVTDROSPPVSPPIQPHFIINQIELYLGSSNNNTLKYSAG 482
Db 336 NTNRLSKQLITLLPGE-FWTHLSINRPHQTIAINKLSLQIKIVFTTFKNDN--EYQKN 392

QY 483 GLSLNY---QMTTFQFPRKDCNLVIDPGCSPNFNNSYSHILSHFSFTYSYVIGLQLOI 539
Db 393 FNVANNQEPQETT-----NYPNDYGGG-NSQKFKHNLSHFPLIIH-----KLEF 435

QY 540 LD-----TGVLGWTHSSVDRYNAISDKIITMIPAIGNNLDTNKSIVIEGPGHTGGNLYVLO 595
Db 436 AEYPHISFALGWTSHSVNSQNLISSESVSQTPLVKAYEV-TNNSVIRGSGFTGGDLIELR 494

QY 596 SQGRLEITCETPNSTQSYFIRLYATNGAGNTLPNLSLTIPGVIGIPPORLANTTSGTNY 655
Db 495 D--KCSIKCKA-SSLKKYAIISLFYAANNAIAVSDIVDGSAGVL-----LOFTSRKGN 545

QY 656 NN-----LOYGDFGYFQFPSTVTLPLNRNIPFIENRAD--VSNSTIIDKIEFIPITSM 708
Db 546 NNFTIQDLNVDKQVHTLLVDIELPESEIHLKREDDYBEGVILLIDKLEFKPIDENY 605

QY 709 HQNREKOKLETIQTINTFTTNHTKTL 735

Db 606 ---TNEMLLEKAKKAVNVLFINATNAL 629

Search completed: December 4, 2005, 13:04:00
Job time : 169 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2005, 12:51:02 ; Search time 11 Seconds
(without alignments)

319.948 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869

Sequence: 1 MQQNDNNEYIIDSHSPY.....KLETIQKINTFFNHTKTL 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 26661 seqs, 4788334 residues

Total number of hits satisfying chosen parameters: 26661

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New.*

- 1: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	703	18.2	1210	7	US-11-058-727-4
2	703	18.2	1210	7	US-11-058-727-4
3	689.5	17.8	1386	7	US-11-091-643-6
4	678.5	17.5	673	7	US-11-058-727-14
5	678.5	17.5	673	7	US-11-058-727-14
6	675	17.4	1206	7	US-11-058-727-2
7	675	17.4	1206	7	US-11-058-727-2
8	669	17.3	1316	7	US-11-091-643-4
9	657.5	17.0	675	7	US-11-058-727-74
10	657.5	17.0	675	7	US-11-058-727-80
11	657.5	17.0	675	7	US-11-058-727-80
12	657.5	17.0	675	7	US-11-058-727-80
13	657	17.0	674	7	US-11-058-727-82
14	657	17.0	674	7	US-11-058-727-82
15	656.5	17.0	675	7	US-11-058-727-42
16	656.5	17.0	675	7	US-11-058-727-48
17	656.5	17.0	675	7	US-11-058-727-48
18	656.5	17.0	675	7	US-11-058-727-48
19	656	17.0	674	7	US-11-058-727-50
20	656	17.0	674	7	US-11-058-727-50
21	654.5	16.9	673	7	US-11-058-727-70
22	654.5	16.9	673	7	US-11-058-727-70
23	653.5	16.9	673	7	US-11-058-727-34
24	653.5	16.9	673	7	US-11-058-727-68
25	653.5	16.9	673	7	US-11-058-727-34

ALIGNMENTS

RESULT 1

US-11-058-727-4
; Sequence 4, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Heirmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Fresnall
; APPLICANT: James P.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1210
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-11-058-727-4

Query Match 18.2%; Score 703; DB 7; Length 1210;

Best Local Similarity 27.8%; Fred.No. 2.5e-46;
Matches 216; Conservative 127; Mismatches 292; Indels 142; Gaps 31;

QY	5	NDNNEYIIDSHSPYFPNRSND-S-RYPYTNPNQPLQNTNYKWLNNCOGN-TOYGDN 62
Db	4	NNQNEYIIDATPS-----TSVNSDNRYPFANEPTNALQNMDYKDYLKMSAGNASYPGS 59
QY	63	FETPASADTIAAVSAGTIIVSGTLLAGIGLTSISGPIGIIGAIISFGYLIITVFWPAGEQ 122
Db	60	FEVLVSGQD--AAKAAIDIVGKLLSLG-----VPPVGFIVSLYTLQIDILWPSGQ- 108
QY	123	DKTWTQFIKNGEIVFDVPLTFESIKQLKLTLEGFRQLQSYNTALDDWRKLRKLAQPL 182
Db	109	-KSQWEIFMEQVEELINQKIAFYARNKALSELEGLGNNYQLYLTALKEWKE----- 158

183 PPSSALQQAALTKIRPENVNDPIREIPGQLETKTLLPIYAQAANFHLNLLQGAE 242
159 NPNGS--RALRDVRNRFEILDSLFTQMPFSFRVTNEVPEPLTVYTOAANLHLLLLKDAASI 216
243 LADEWNADIHPSQIBENAGTSDYYKLLKENIPKYSNYCANTYRTGLKNLRDEPNMKWSI 302
217 FGEWGW-----STTTINNYDRQMKLTAEYSDHCVKWYETGLAKUGTSKQWVD 267
303 FNDYRYMTITVLTDSIQSLDIKRYRDSIGIEVKIGKNELTREIYTFEINFDRLPOL 362
268 YNQFRREMTLVLDVVALFPNYDTRY-----PMETKA---OLTREYVT-----DPLGAV 314
363 RVQ-----PNLATMEYNLTASFKLFSLEQIFYYTENTNFGN-----RLVGISNR 408
315 NVSSIGSWYDKAPSGFVIESSVIRPP-HVFDYITGLTVYVTSQSRSSISSARYIRHWAGHQS 373
409 DAPTYNTITETLYGERTGSPTKTIRPPESYKVSIVTDROQSPVPIQP-----HFIIN 463
374 YHRIFSDNIHKOMYGTNQNLHSTSF-DFTNYDIYKTLSDAVLLDIVPPGYTYIFFGMP 432
464 QIELYLNGSSNN---TLKYS-----AGGSLSNYQNTTFFQPPRKKDCNLVIDPGCS-- 511
433 EVEFFWVQNLNTRKTLKYNPVSCKDIAG-----TRDSELELPETSQ 476
512 PNFNNYSHLSHF-SLFTYSYVIGLQILDGTGLGWTHSSVDRYNAISDKIITMIPAIC 570
477 PNYESYSHRLCHITSIPATGSTTGL-----VPVFSWTHRSADLINAVHSDKITQIPVVK 530
571 GNNL-----DTNSKVIIEGPGHTGNNLVYLSQSG-----RLBITCETPNSTQSYFIRLR 618
531 VSDLAFTSGPNTVSGFVIESSVIRPP-HVFDYITGLTVYVTSQSRSSISSARYIRHWAGHQS 586
619 YATNGAGNTLPNLSLTIPGVIIPQORLNTFTSGTNNYNNLYQDGFYGFQFPPSTVTLPLNR 678
587 YAS--ANNTEFYINPSEENVKS-HAQTWNRGKALTYNKNFYATLPPIKFTTTE----- 637
679 NIPF-----IFNRADVSNLSIIDKIEFIPITSSMHQNRKQKLETIQTINTFTFN 730
638 --PFITLGAIFEAEDFLGIEAYIDRIEFIPVDETY---EAEQDLEAAKAVNALFTN 689

RESULT 2

US-11-108-389-4
; Sequence 4, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F. H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 357118/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1210
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis

US-11-108-389-4

Query Match 18.2%; Score 703; DB 7; Length 1210;
Best Local Similarity 27.8%; Pred. No. 2.5e-46;
Matches 216; Conservative 127; Mismatches 292; Indels 142; Gaps 31;
QY 5 NDNNEVHIIDSHTSPPYPPNRNSNDS-RYPYTNPNQPLONTNYKWLNMCOGN-TQYGDN 62
DB 4 NNQNEYIIDATPS---TSVSNDSNRYFPANEPNTALQNNMDYKDYKMSAGNASEYPGS 59
QY 63 FETPASADTTAAVSAGTIVSGTLLAGGLTISISGPIGIIGAILISFGTLLTVFWPAGEQ 122
DB 60 PEVLVSGQD--AKAAIDIVGKLSGLG-----VPFVGPIVSLYLTQILDILWPSQ-- 108
QY 123 DKTWVTQFKMGBIFVDVPTLSTESIKQLKLTLEGFRQILQSYNTALDDWRKLKRLQAPGL 182
DB 109 -KQWEIFMGEVEELINQKIAEYARNKALSELEGLGNNYQLYLTALBEWKE----- 158
QY 183 PPSALQQAALTKIRPENVNDPIREIPGQLETKTLLPIYAQAANFHLNLLQGAE 242
DB 159 NPNGS--RALRDVRNRFEILDSLFTQMPFSFRVTNEVPEPLTVYTOAANLHLLLLKDAASI 216
QY 243 LADEWNADIHPSQIBENAGTSDYYKLLKENIPKYSNYCANTYRTGLKNLRDEPNMKWSI 302
DB 217 FGEWGW-----STTTINNYDRQMKLTAEYSDHCVKWYETGLAKUGTSKQWVD 267
QY 303 FNDYRYMTITVLTDSIQSLDIKRYRDSIGIEVKIGKNELTREIYTFEINFDRLPOL 362
DB 268 YNQFRREMTLVLDVVALFPNYDTRY-----PMETKA---OLTREYVT-----DPLGAV 314
QY 363 RVQ-----PNLATMEYNLTASFKLFSLEQIFYYTENTNFGN-----RLVGISNR 408
DB 315 NVSSIGSWYDKAPSGFVIESSVIRPP-HVFDYITGLTVYVTSQSRSSISSARYIRHWAGHQS 373
QY 409 DAPTYNTITETLYGERTGSPTKTIRPPESYKVSIVTDROQSPVPIQP-----HFIIN 463
DB 374 YHRIFSDNIHKOMYGTNQNLHSTSF-DFTNYDIYKTLSDAVLLDIVPPGYTYIFFGMP 432
QY 464 QIELYLNGSSNN---TLKYS-----AGGSLSNYQNTTFFQPPRKKDCNLVIDPGCS-- 511
DB 433 EVEFFWVQNLNTRKTLKYNPVSCKDIAG-----TRDSELELPETSQ 476
QY 512 PNFNNYSHLSHF-SLFTYSYVIGLQILDGTGLGWTHSSVDRYNAISDKIITMIPAIC 570
DB 477 PNYESYSHRLCHITSIPATGSTTGL-----VPVFSWTHRSADLINAVHSDKITQIPVVK 530
QY 571 GNNL-----DTNSKVIIEGPGHTGNNLVYLSQSG-----RLBITCETPNSTQSYFIRLR 618
DB 531 VSDLAFTSGPNTVSGFVIESSVIRPP-HVFDYITGLTVYVTSQSRSSISSARYIRHWAGHQS 586
QY 619 YATNGAGNTLPNLSLTIPGVIIPQORLNTFTSGTNNYNNLYQDGFYGFQFPPSTVTLPLNR 678
DB 587 YAS--ANNTEFYINPSEENVKS-HAQTWNRGKALTYNKNFYATLPPIKFTTTE----- 637
QY 679 NIPF-----IFNRADVSNLSIIDKIEFIPITSSMHQNRKQKLETIQTINTFTFN 730
DB 638 --PFITLGAIFEAEDFLGIEAYIDRIEFIPVDETY---EAEQDLEAAKAVNALFTN 689

RESULT 3

US-11-091-643-6
; Sequence 6, Application US/11091643
; Publication No. US20050246789A1
; GENERAL INFORMATION:
; APPLICANT: TANAKA, Masao
; APPLICANT: YOKOYAMA, Tomoko
; APPLICANT: AOYAGI, Moriichi
; APPLICANT: HASEGAWA, Makoto
; APPLICANT: EHARA, Gaku
; APPLICANT: KIMURA, Masaharu
; APPLICANT: NISHIHASHI, Hideji
; TITLE OF INVENTION: Polypeptide having larvae growth inhibiting or
; TITLE OF INVENTION: insecticidal effect on scarabaeidae insects and

```
; TITLE OF INVENTION: poynucleotide encoding the same
; FILE REFERENCE: Opl335
; CURRENT APPLICATION NUMBER: US/11/091,643
; CURRENT FILING DATE: 2005-03-29
; PRIOR APPLICATION NUMBER: JP 2001-115754
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: JP 2001-203463
; PRIOR FILING DATE: 2001-07-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1386
; TYPE: PRT
; ORGANISM: Bacillus popilliae
; US-11-091-643-6

Query Match      17.8%; Score 689.5; DB 7; Length 1386;
Best Local Similarity 30.6%; Pred. No. 3.3e-45;
Matches 239; Conservative 113; Mismatches 286; Indels 143; Gaps 38;

Qy 1 MNQ---NNDN-----NEYEII--DSHTSPYFNRSNDSRYPTNNPQLONTNYKE 48
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19 MNQHYNQNDKSNQSGNEVQIIQFSSNALLYSPN-----KYPYATDPNVIAEGRSYKN 72

Qy 49 WLMNCOQNTQYGNFETFPASADTTAAVSAGTIVSGTLLAGIGGLTS--ISGPIGI--ICAI 105
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73 WLMCMVG---VGDD-----TRSPEARVTAOSSISTSL-----GITSTIIGALGIPVVGEEA 119

Qy 106 IISFGTLITVFWPAGEQDKTWTQIFKMGEIFVDTPTTESIKQLKQTLQLEGFROILOSYN 165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 120 IGIFGALLDMLWPAGADP---WVIFPMHVEELINSKITETVKNCAITRLDGLGNVLAIYQ 176

Qy 166 TALDDWRKLRQLAPGPSALQQAALTUKIRFENVHDFIREIPGQLEYTKYTKLLPI 225
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 177 KAPEWQ---QHP-----TLSEARLRVTDDFSNVKNKPEAFMPSPRVGVEVPLLSV 225

Qy 226 YQAQANFHLNLOQAELEADENWADIHPSQIEPNAGTSDDYKLLKENIKPKSYNCANTY 285
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 226 YVSAANLHLLLRDSSIFGLDWGL-----SOTHVN-----DNYNLQIRRSADYANHCTTWY 276

Qy 286 RTGLKNLRDEPNMKWSIFNDYRYMTITVLDTISQFSLYDIKRYRDSIGGIEVGKINEL 345
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 277 RTGLRLOQTNASWVWYNFRREMTITVLDVCLFSSYDRSY-----PMELRG---EL 328

Qy 346 TREIYITEINFDRLPQLRQVONLMEYNLTRASFKLFSLEQFIPTENTNFGNRLVGI 405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 329 TREIYTDVCGASP--WYNRAPNPASIENTVVRQPHP--FTWLVTITVNTGQVRSQD---GN 382

Qy 406 SNRDAPYSNTITLTYGERTGSP-----TKTIRPFESYKVSIVTDRO---SPVSPIQPHF 460
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 383 SNYVWKSHTQVSETGSGGPIQSPCTGSGTGI-----YR-----TDNLLFNP-----F 425

Qy 461 IINQIEL-----YLNSSNNLTLYSAGSLSN-----YQNTTFQFPKKDCNVIDPCCS 511
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 426 LLGDIYTINGYSYLANLFGIYSARPTTKRSIELLYENORVFPAYNHQIREL---PGVN 482

Qy 512 ---PNFNYSHILSHFLTYVYVIGLQLQILDGTG---VLGWTSHSSVDRYNALIS-DKII 563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 483 SDRPTAADYSHRSLSYISGFA-----TDVGGTVLVYGVTSSTATRENNITLDDRI 531

Qy 564 TMIPAIGNNLDNLSKVIPEGHGTGNLVLQSQGRLEITCETPNSTQSYFIRLYATNG 623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 532 VOLPAVKGTSLN-NCQVVRGTGFTGGDWLKNNGNFTSLALGF-RSTVTVRLRIRYAAAA 589

Qy 624 AGNLTPLNISUTIPGVIPIPORLNNNTSGTNYN---NLQYGDGFGYFQPPSTVTI----- 674
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 590 GGS---GFSLVISDQYGEFF--TTTVSLSSSTMYSLPQNVVPYEAFFKIVDLPSVTIRNTSPA 645

Qy 675 --PLNRNIPFENRADVSNSTLIIDKIEFIPITSSMHQNRKQKLETIQTINTFFTNHT 732
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 646 STTFRDLFRFIVPLIGILAN--ILIDRIEFVPIEGSLFEYETKQKLEKARKAVNHLFTDGS 703

Qy 733 K 733
```

Db 704 K 704

RESULT 4

```
US-11-058-727-14
; Sequence 14, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 673
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis (truncated)
; US-11-058-727-14
```

Query Match 17.5%; Score 678.5; DB 7; Length 673;

Best Local Similarity 27.8%; Pred. No. 8e-45;

Matches 209; Conservative 123; Mismatches 280; Indels 139; Gaps 30;

```
Qy 5 NDNNEYEIIIDSHTSYPFPNRSNDS--RYPYTNPNQPLQNTNYKELNMCQGN--TQYGDN 62
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4 NNQNEYEIIIDTAPS-----TSVSNDSNRYFPANEPTNALQNMDYKDYLRKMSAGNASYPSGS 59

Qy 63 FETPASADTTAAVSAGTIVSGTLLAGIGGLTSISGPIGIIGAIISFGFLITVFWPAGEQ 122
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 60 PEVLVSGQD--AAKAAIDIVGKLLSGLG-----VPFVGPIVSLYTLQILDILWPSGQ- 108

Qy 123 DKTVMTOFIRMGEIFVDTPTTESIKQLKQTLQLEGFROILOSYN TALDDWRKLRQLAPGL 182
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109 -KSQWEIFMEQVEELINQIAEYARNKALSELGLGNVQLYLTALKEWKE----- 158

Qy 183 PPSALQQAALTUKIRFENVHDFIREIPGQLEYTKYTKLLLPYQAQANFHLNLOQGA 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 159 NPNGS--RALRDVRRNRPEILDSLEFTQYMPSPRVNTNPFVFPFLTVYTTQAANLHLLLKDASI 216

Qy 243 LADSNADIHPSQIEPNAGTSDDYKLLKENIKPKSYNCANTYRTGLKNLRDEPNMKWSI 302
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 217 FGEWGW-----STTTNNYDROMKLTAEYSDHCVKWYETGLAKLGTSAKQWVD 267

Qy 303 FNDYRRYMTITVLDTISQFSLYDIKRYRDSIGGIEVGKIKNELTREIYITEINFDRLPOL 362
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 268 YNQFRREMTITVLDVVALFPNYDTITY-----PMETKA---QLTREYVT-----DPLGAV 314

Qy 363 RVQ-----PNLATMEYNLTRASFKLFSLEQFIPTENTNFGN-----RLVGISNR 408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 315 NVSSIGSWYDKAPSGFVIESVIRPP-HVFDYITGLTVYTTQSSRSISSARYIRHWAGHOIS 373

Qy 409 DAPYSNTITLTYGERTGSP-----TKTIRPFESYKVSIVTDROSPVPSPIOP-----HFIIN 463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 374 YHRIFSDNIIKOMYGTNQNLHSTSTF-DFTNYDYIKTLSKDAVLLDIVFPYGYTIFFGMP 432
```

QY 464 QIELYNGSSNN---TLKYS-----AGSLSNYQNTTFFQPRKDCNVLVIDPGCS-- 511
Db 433 EVEFFMVNQLNTRKTLKYNPVSQDIAG-----TRDSELEPPETSDQ 476
QY 512 PNFNNYSHLISHF-SLFTYSYVIGLQILDTGVLGTHSSVDRYNAISDKIITMIPA 570
Db 477 PNYESYSHRLCHITSIPATGTTGL-----VPVFSWTHRSADLINAVHSDKITQIPVVK 530
QY 571 GNNL-----DTNSKVIPEGHTGNNLVYLSQSG-----RLBITCETPNSTQSYFIRLR 618
Db 531 VSDLAPISITGGPNNTVSGPFTGGGIIKVRNGVISHMRVKIS-----DINKEYSMRIR 586
QY 619 YATNGAGNTLPNLSLTIPGVIGIPQRLNNTFSGTNNYNLQYDGFQFPQSTVTLPLNR 678
Db 587 YAS--ANNTEFYINPSEENVKS-HAOKTMNRGEALTYNKFNYATLPPIKFTTTE----- 637
QY 679 NIPF-----IFNRADVNSNLIIDKIEFIPI 704
Db 638 --PFITLGAIFAEDFLGIEAYIDRIEFIPV 666
RESULT 5
US-11-108-389-14
; Sequence 14, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE REFERENCE: 357118/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; TYPE: PRT
; LENGTH: 673
; ORGANISM: Bacillus thuringiensis (truncated)
US-11-108-389-14
Query Match 17.5%; Score 678.5; DB 7; Length 673;
Best Local Similarity 27.8%; Pred. No. 8e-45;
Matches 209; Conservative 123; Mismatches 280; Indels 139; Gaps 30;
QY 5 5 NNONEVEIIDSHTSPFPNRSNDS-RYPYTNPNQPLQNTYKWLNMCCGN-TQYGDN 62
Db 4 NNONEVEIIDSHTSPFPNRSNDS-RYPYTNPNQPLQNTYKWLNMCCGN-TQYGDN 62
QY 63 FETPASADTIAVSACTIVSGTLLAGIGLTSISGPIGIIGAILISFGTLITVFPAGEQ 122
Db 60 PEVLVSGQD--AKAAIDIVGKLLSLG-----VPFVGPIVSIVTQIDILVPSGQ- 108
QY 123 DKTVMTOFKMGSEIFVDTPLETESIKQLQTLQLEGFQRQILQSYNTALDDMRKRLQAPGL 182
Db 109 -KSQWEIFMEQVEELINQKIAEYARNKALSELGLNNGYLYLTAEWKE----- 158
QY 183 PPSALQQAALTKIRFENVNDHFIREFGQFQETKTLPIYQAANFHLNLLQGA 242
Db 159 NPNGS--RALRDVRNRFEILDSLFTQMPFSFRVTNFEVPELTVYTAANLHLLKLDASI 216

QY 243 LADENWADIHPSQIEPNAGTSDDDYYKLLKENIPKYSNYCANTYRTGLKNLDEPNMKWSI 302
Db 217 FGBEWG-----STTTINNYDRQMKLTAEYSCHVKWYETGLAKGTSAKQWD 267
QY 303 FNDYRRYMTITVLDTSIQFSLYDIKRYRDSIGGLEVKIKNELTREIYTTINPRLPOL 362
Db 268 YNOFRREMTITVLDVVALFPNYDTRTY-----PMETKA--OLTREVTY-----DPLGAV 314
QY 363 RVQ-----PNLATMEYNLTRASPKLFSFLQFIPYTYTENTNFGN-----RLVGISNR 408
Db 315 NVSSIGSWYDKAPSGVIESVIRPP-HVFDYITGLTYTQSRSSISSARYIRHWAGHQS 373
QY 409 DAPYSNITITETLYGERTGSPTKTIRPFSYKYSIVTDROSPPVSPQIP-----HFIIN 463
Db 374 YHRIFSDNIIKQMYGTNQLHSTF-DFTNYDIYKTLSDAVLLDIVFGVYTYFFGMP 432
QY 464 QIELYNGSSNN---TLKYS-----AGSLSNYQNTTFFQPRKDCNVLVIDPGCS-- 511
Db 433 EVEFFMVNQLNTRKTLKYNPVSQDIAG-----TRDSELEPPETSDQ 476
QY 512 PNFNNYSHLISHF-SLFTYSYVIGLQILDTGVLGTHSSVDRYNAISDKIITMIPA 570
Db 477 PNYESYSHRLCHITSIPATGTTGL-----VPVFSWTHRSADLINAVHSDKITQIPVVK 530
QY 571 GNNL-----DTNSKVIPEGHTGNNLVYLSQSG-----RLBITCETPNSTQSYFIRLR 618
Db 531 VSDLAPISITGGPNNTVSGPFTGGGIIKVRNGVISHMRVKIS-----DINKEYSMRIR 586
QY 619 YATNGAGNTLPNLSLTIPGVIGIPQRLNNTFSGTNNYNLQYDGFQFPQSTVTLPLNR 678
Db 587 YAS--ANNTEFYINPSEENVKS-HAOKTMNRGEALTYNKFNYATLPPIKFTTTE----- 637
QY 679 NIPF-----IFNRADVNSNLIIDKIEFIPI 704
Db 638 --PFITLGAIFAEDFLGIEAYIDRIEFIPV 666
RESULT 6
US-11-058-727-2
; Sequence 2, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE REFERENCE: 357118/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1206
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-11-058-727-2
Query Match 17.4%; Score 675; DB 7; Length 1206;
Best Local Similarity 28.2%; Pred. No. 3.5e-44;

	Matches	216;	Conservative	125;	Mismatches	300;	Indels	126;	Gaps	32;
Qy	5	NDNNEYIIDSHTSPYFNNRNSDS--RYPYTNNPNQLONTNYKEWLNMCQN--TQYGDN	62							
Db	4	NNQNEYIIDATPS-----TSVSDNSNRYPANEPNALQNMWDYKDYLKMSAGNASASEPGS	59							
Qy	63	PETPASADTTAAVSAGTIVSGTLIAGIGGLTSISGPIGIIICALLISFGTLITVPWPAGEQ	122							
Db	60	PEVLVSGQD--AKAAADIDVGKLLSGLG-----VPFVGPIVSLVTQLDILWPSGE-	108							
Qy	123	DKTVTTOFIKWGEIFVDTPLTTESTIKQLKTLEGFRQILOSYNTALDDWRKLKLQAPGL	182							
Db	109	-KSQWEIIMEQVEELINQKIAEYARNKALSELEGGNNYQLYLTALEEWE-	158							
Qy	183	PPSSALOQAALTAKIRPENVHNDIREIPGOLETYKTLILLPIYAQAANPHNLNQOAE	242							
Db	159	NPNGS--RALDVDNRREILDLSFTQMPSRVTNFEVFFLVTVYAMAANHLHLKKDASI	216							
Qy	243	LADENNADIHPQSIEPNAGTDDYYKLLKENIPKYSNYCANTYRTGLKNLDEPNMKWSI	302							
Db	217	FGEHWG-----STTINNYYDRQMKLTAEYSCHVKWYETGLAKLGTSAKQWVD	267							
Qy	303	FNDVRRYMTITVLDTISQFSLYDIKRYRDSIGGIEVRGKIKNELTREITYTTEINFRLPQL	362							
Db	268	YNQPRREWTLAVLDVALFPNYDRTY-----PMETKA---QLTREYVT-----DPLGAV	314							
Qy	363	RVO-----PNLATWEYNLTRASFKLSFLSQFIPTENTNFGN-RLV-----GIS	406							
Db	315	NVSSIGSWYDKAPSFGVIESVIRPP-HVFHYITGLTYTQSRSISSARYIRHWAGHQIS	373							
Qy	407	NRDAPTYSNTITETLYGERTGSPTKIRPPESKVSVITDRQPPVSPTOP-----HFI	461							
Db	374	YHRVRSRGSN--LOOMYGTONQLHSTSTF-DFTNYDVIYKLSKDVAVLDDIVPGYTIYFFG	430							
Qy	462	INOELYLNGSSNN--TLKYSAGGSLSNYQNTTFFQFPRKDCQNLVIDPCGCSFNENYS	518							
Db	431	MPEVEFFWNQNLNTRKTLKY-----NPVSKDIIASTRDESEULELPETSQQNYSESYS	483							
Qy	519	HILSHF-SLFYTYSYVIGIQILDTGVLGWTHSSHVDRYNAISDXIIITWIPAik-GNNLDT	576							
Db	484	HLCLHTSIPATGNWTGL-----VPFVSWTHRSDLNNITYSKDIQTQIPAVKCWDNLPP	537							
Qy	577	NSKVIEGPGHTGGLNV-YLOSQG-----RLBITCTPNSTOSYFIRLRYATNGAGNT	627							
Db	538	-VPVVKGPGHTGGDLQVNRSTGVSGLTLFARYGLEA---KAGKYRVRURYATDA----	589							
Qy	628	LPNISLTIPGVI GPIORLNNFTSGTWNVNLOYGDGFQPPSTVTIPLNKRNIPIFNRA	687							
Db	590	--DIVLHVNDAQIQMPKTMNPG-----EDLTSKTFKVADAITTLNLTADSSLALKHNLG	641							
Qy	688	DVNSN-----ILIIDKIEFIPITSSHHQNRKQKETIQTKINTFTTN	730							
Db	642	EDPNSTLSGIYVDRIEFIPVDETY--EAEQDLEAAKAVNALFTN	685							

RESULT 7

```

US-11-108-389-2
; Sequence 2, Application US/11108389
; Publication NO. US20050261180A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108.389

```

```

? CURRENT FILING DATE: 2005-04-18
? PRIOR APPLICATION NUMBER: 60/391,786
? PRIOR FILING DATE: 2002-06-26
? PRIOR APPLICATION NUMBER: 60/460,787
? PRIOR FILING DATE: 2003-04-04
? PRIOR APPLICATION NUMBER: 10/606,320
? PRIOR FILING DATE: 2003-06-25
? NUMBER OF SEQ ID NOS: 134
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 2
? LENGTH: 1206
? TYPE: PrT
? ORGANISM: Bacillus thuringiensis
US-11-108-389-2

```

Query Match	17.4%; Score 675; DB 7; Length 1206;
Best Local Similarity	28.2%; Pred. No. 3.5e-44;
Matches	216; Conservative 125; Mismatches 300; Indels 126; Gaps 32;
QY	5 NDNNEYEIIDSHTSYPPFNNSNDS--RYPVTNNPNQPLQNTYKWEMLNMQGN--TOYGDN 62
DB	4 NNQNEVEIIDATPS-----TSVNSDSNRYPANETNALQNDYKYDKLMSAGNASEYPGS 59
QY	63 FETTFASADTTAAVSAGTIVSGTLLAGIGGUTSISGPTIGIAGIIISFGTILITVFWPAGEQ 122
DB	60 PEVLVSQD--AAKAAIDIVKLLSGLG-----VPFVGPIVSLYQTQLIDILWPSGE- 108
QY	123 DKTVWTQPIKMGELFVDVTPLTESI KQLKLTQTSBGFQRIQLOSNTFALDDWRKLKLQAPGL 182
DB	109 -KSGWEIFMEQVEELINQKIAEYARNKAUSELGLGNYYQLYTALBEEF----- 158
QY	183 PPSALQQAALTKIRFENVHNDFI REIPGFOLETYKTLTLLPIYAQAANFHLNLLQOGAE 242
DB	159 NPNGS--RALRDVNRFEILDSLFTQYMPFSRVTPNEVPFLTYYAMAANLHLLLLKDASI 216
QY	243 LADEWNADIHPSQIEPNAGTSDDYKLLKENIIPKYSNYCANTVRTGUKNLRDEPNMKWSI 302
DB	217 FGBEWG-----STTTINNYDQMKLTAEYSDHCWKVYETGLAKLKGTSAKQWVD 267
QY	303 FNDYRRVMTITVLDTTSQFSLYDKIRYRDSIGGIEVKGIKNELTRELYTTEINFDRPLQL 362
DB	268 YNFRFRMTLAVLDVVALFPNYDRTY-----PMETKA-----QLTREYVT-----DPLGAV 314
QY	363 RVQ-----ENLATMEYNLTRASFKLFSLEQEIFYFTENTNFGN-RLV-----GIS 406
DB	315 NVSSIGSWYDKAPSGFVIESSVIRPP-HVPDYITGLTVYTSQRSISSARVIRHWAGHQS 373
QY	407 NRDAPYSNTIITELGVERTGSPTTKIRPFESYKVISIVTDROSPPVSPTOP-----HFI 461
DB	374 YHRVSRGSN--LQOMYGTGNQLHSTSTP-DFTNYDIYKTLSDKAVALDIDVYPGYTYIFFG 430
QY	462 INQIELVYLGSSNN---TLKYSAGGSLSNYQNTTFFQFPRKDCNLVIDPGCSFNFNYS 518
DB	431 MPEVFEFVWQNLNTRKTLKY-----NPVSKDITASTRDSLELPPETSDQDNYESYS 483
QY	519 HILSHF-SLFTYSYVIGLQILDTGVLGWTHSSVDRYNAISDKIITMIPA-KGNNLDT 576
DB	484 HRLCHITSIPATGNTTGL-----VPVFSWTHRSADLNNTIYSDKITQIPAVKCDWNLPP 537
QY	577 NSKVIKPGHGTGNLV-YLOSQ-----RLBITCTETPNSTQSVIRLRYATNGAGNT 627
DB	538 -VPVWKGPHGTGDLQYKNRSTSGVTLFLARYGLALE---KAGKYRVRURYATDA---- 589
QY	628 LPMISLITPGVIGIPQRLNNTPSGTYNNLQYDGFYFQFPSTVTTLPLNRNPFFIENRA 687
DB	590 --DIVLHVNDQAQIQMPKTNPG-----EDLTSTKFKVADAITTLNLATDSSLAKHNLG 641
QY	688 DVNSG-----ILLIDKTEFPITSSMQNRKQKLETIQTINKTPFTN 730
DB	642 EDPNSTLSGVYVDRIEFIVDETY---EAEQDLEAAKAVNALFTN 695

RESULT 8

```
US-11-091-643-4
; Sequence 4, Application US/11091643
; Publication No. US20050246789A1
; GENERAL INFORMATION:
; APPLICANT: TANAKA, Masao
; APPLICANT: YOKOYAMA, Tomoko
; APPLICANT: AOYAGI, Morichi
; APPLICANT: HASEGAWA, Makoto
; APPLICANT: EHARA, Gaku
; APPLICANT: KIMURA, Masaharu
; APPLICANT: NISHIHASHI, Hideji
; TITLE OF INVENTION: Polypeptide having larval growth inhibiting or
; TITLE OF INVENTION: insecticidal effect on scarabaeidae insects and
; TITLE OF INVENTION: polynucleotide encoding the same
; FILE REFERENCE: OP1335
; CURRENT APPLICATION NUMBER: US/11/091,643
; CURRENT FILING DATE: 2005-03-29
; PRIOR APPLICATION NUMBER: JP 2001-115754
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: JP 2001-203463
; PRIOR FILING DATE: 2001-07-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1316
; TYPE: PRT
; ORGANISM: Bacillus popilliae
US-11-091-643-4

Query Match 17.3%; Score 669; DB 7; Length 1316;
Best Local Similarity 29.8%; Pred. No. 1.2e-43;
Matches 230; Conservative 106; Mismatches 301; Indels 136; Gaps 34;

QY 1 MNQ---NNDN-----NEVEII--DSHTSPYPNNSNDSRYPTNNPNQPLONTNYKE 48
Db 19 MNQYHNDKNSYNQSGNEWIIQPSNALLYSPN-----KPYATDPNVIAEGSYKN 72
QY 49 WLMNCOGNTQYDGNFETASADTIAAVSAGTIVSGTLLAGIGLTSIGSPIG-----IG 103
Db 73 WLDNCTGT---GDTRSPETAISKGAVSAAITIS-----TGLLGLLVPPASQIG 119
QY 104 AIIISFGTLITVFWPAGEQDKTWTQPIKMGELFVDTPLTESI KQLKLTQLEGFROILOS 163
Db 120 AF---YTFLLNTLWPA---SNTQWEOFIARVEELINAKLTDHVRNSALTKLNGLRNIEI 173
QY 164 YNTALDDWEKLRLOAPGLPPSSALQQAALTAKIRENVHNDFIRIPGQLETYKTLIL 223
Db 174 YNEALIVKQ-----DPNNSKLKD---DVRSKFVGLNSQFEYIPOFKERGFEVQLL 222
QY 224 PIYAQAANFHLNLLQQAELADEWNADIHPSQIEPNAGTSDDYKLLKENIPKSYNCAN 283
Db 223 TIYAQSANLHLLLRDSSLYGASWGF-----AQATIDNNYNQIRKTAEYANHCTT 273
QY 284 TYRTGLKNLRDBENMKWSIFNDYRRYMTITVLTISQFSLYDIKRYRDSIGGIEVKIGKN 343
Db 274 WYQTLQRLQGTASSWLSYHFRREMTITVLDICALFSNYDARSY-----PLEVRG--- 325
QY 344 ELTREIYTYTEIN-----FRLPOLRQPNLA-TMEYNLTFAASKFLSFLEQFIYFNTN 397
Db 326 ELTREIYTDVAGTWNIDRAPSFABIELNIVRAPRTVWISGDLIVYGRLYGYGTGND 385
QY 398 F--GNRLVGISNR---DAPTYNTITETLYGERTGSPPTKTRPFESYKVSIVTDQSP 451
Db 386 YNAHRLDPLETNGRYFEGTYGSTINIS-----RTUSIPWNSI---DVSSTVTVTGSW 438
QY 452 PVSPIQPHFIINQIELYLNGSNNTLYKSAGGSLSN--YQNTTFFQPRKOCNCLVID-P 508
Db 439 PTG----GFVLGVASARFFSKSPST-----GLLGERVQNPVYF-----SSSTLTFLNP 483
QY 509 GC---SPNFNNYSHILSHSLFTYSYVIGLQILDTGLVWTHSSVDYRNALSKLITM 565
Db 484 GVDQDPTTAADYSKLSCTITAPR---TGINGTV---PVFGYSATVSRDNRLEPKITQ 536

US-11-058-727-74
; Sequence 74, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE REFERENCES: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis (mutated)
US-11-058-727-74

Query Match 17.0%; Score 657.5; DB 7; Length 675;
Best Local Similarity 28.5%; Pred. No. 3.3e-43;
Matches 211; Conservative 120; Mismatches 293; Indels 117; Gaps 31;

QY 5 NDNNEYIIDSHTSPYFPNNSNDS-RYPYTNPNQPLQNTYKELNMCQGN-TOYGDN 62
Db 4 NNQVEIIDATFS-----TSVSNDSNRYPANEPFNALQNDYKDYLMKSGNASEVPGS 59
QY 63 FETFASADTIAAVSAGTIVSGTLLAGIGLTSIGSPIGIIIGAILISFGTLITVFWPAGEQ 122
Db 60 PEVLVSQD--AKAAIDIVGKLLSGLG-----VPFVGPIVSLYTLQIDILWPSCG- 108
QY 123 DKTWTQPIKMGELFVDTPLTESI KQLKLTQLEGFROILOSNTALDDWRKLRLOAPGL 182
Db 109 -KSQWEIFMEQVEELINQKIAEYARNKALSELEGNNYQLYLTALBEEWEE-NPFRSRGF 166
QY 183 PPSALQQAALTAKIRENVHNDFIRIPGQLETYKTLILPIYAQAANFHLNLLQQAEL 242
Db 167 RSGALRD-----VRNRPFLDLSLFTQYMPSRFVTVNPFVTVYMAANLHLLLKQASI 222
QY 243 LADEWNADIHPSQIEPNAGTSDDYKLLKENIPKSYNCANTYRTGLKNLRDBENMKWSI 302
Db 223 FGEWGW-----STTTINNYDRQMKLTAEYSCHVKVYETGLAKLKGTSAKQWVD 273
QY 303 FNDYRRYMTITVLTISQFSLYDIKRYRDSIGGIEVKIGKNELTREIYTYTEINFDRLPQL 362
Db 274 YNQFRREMTILAVLDVVALFPNYDTRY-----PMETKA-----QLTREYVT-----DPLGAV 320
```

Qy	363	RVQ-----P	NLATWYNL	TRASF	KLFL	EQLE	FIFT	TNTN	FGN	RLV-----	GIS	406																																			
Db	321	NVSSIG	WYDKAP	SGVIE	SSVIRP	-HVFD	YITGL	TVTQ	SR	SISSARY	IRHWAG	HOIS 379																																			
Qy	407	NRDAPT	YSNTIT	ETLY	CGERT	SPTK	TI	RP	SE	YKVSV	IVDR	QSPVP	PIQP-----HFI 461																																		
Db	380	YHRV	RSRGN	-LQO	MYGN	QNLH	STF	-DFT	NYDI	YKTL	SKDA	VLLDIV	VEGYT	IFPG 436																																	
Qy	462	INQEL	YLVNG	SSN--	-FLKY	SAG	SLSN	YQNT	TF	PPR	KKDC	NLVDP	GCSP	PNPNYS 518																																	
Db	437	MPEV	EFVW	QNLN	TRK	TLKY	-----	NPV	SKD	IIA	STR	SE	LE	LP	P	TS	DQ	P	NYE	SY	S 489																										
Qy	519	HILSH	-SLFT	YSY	VIGI	QLQ	LD	TG	VLG	WTH	SS	VDR	YNA	ISD	KIT	IM	PAL	K-GNN	LDT 576																												
Db	490	HLR	CHI	TSI	PAT	GNT	TGL-----	VPV	FS	WTH	R	SAD	LNN	TI	YSD	KITQ	IP	PA	VK	CDN	L	PLF 543																									
Qy	577	NSK	VI	EG	PG	HT	G	GN	L	V-YL	OSQ-----	RLE	IT	CT	PN	TS	Q	S	Y	F	IR	L	YAT	NG	AGNT 627																						
Db	544	-VPV	KG	PG	HT	G	DL	LQ	YNR	TS	G	SV	GL	FL	ARY	GL	AL	---	KAG	KY	R	L	R	V	L	AT	DA --- 595																				
Qy	628	LPN	IS	L	T	IP	G	V	I	G	I	P	P	Q	BL	NN	T	FG	S	T	NN	Y	NI	LQ	Y	G	D	F	G	F	Y	P	P	S	T	V	T	L	PL	NR	N	I	P	IF	NR	A 687	
Db	596	--DIV	L	H	V	N	D	A	Q	I	O	M	P	K	T	M	P	G-----	ED	L	T	S	K	T	F	K	V	A	D	A	I	T	T	V	N	L	T	D	S	S	V	A	K	H	N	L	G 647
Qy	688	DV	SN	----	IL	I	D	K	I	E	P	I	PI 704																																		
Db	648	ED	PN	ST	LS	G	I	S	G	I	V	V	D	R	I	E	I	P	V 668																												

RESULT 10

```

US-11-058-727-80
; Sequence 80, Application US/11058727
; Publication No. US20050261483A1
;
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Ciao-Guo Yu
;
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
;
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis (mutated)
US-11-058-727-80

```

[illegible]

Db	60	PEVLVSGQD---AAKAAIDIVGKLSGLG-----VPFVGPIVSLYITQILDILWPSGB-	100
Qy	123	DKTVMTOPIKMGEIFVDTPLTESIQLKQLQTLLEGFRQILQSNTALDDWRKLRQLQAPGL	182
Db	109	-KSQWEIFMEQVEEELINGKIAEYARNKALSELEGLNQYQLVLTAAEWEW-NPFRSRGP	166
Qy	183	PPSALOQAALTAKIRPENVDNFIREFPGOLEYTKTLLPIYAAQANFHLNLLOQGAE	242
Db	167	NGSALARD---VNRNRFELDSLFTQYMPSPFRVTFNFEVFLTVYAMAANLHLLUKDASI	222
Qy	243	LADBNADIHPSQIEPNAGTSDVYVKKLENIPKYSNYCANTRYTGLKNLRDEPNMKWSI	302
Db	223	PGEWGW-----STTTINNYDROMKLTAEYSDHCVKYETGLAKLKTSAKQWVD	273
Qy	303	FNDYRRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKGIKNELTREIYTYTEINFDRPOL	362
Db	274	YNQFRRENTLAVDLVVALFPNYDRTY-----PMETKA---QLTREVYT---DPLGAV	320
Qy	363	RVQ-----PNLATWEYNLTRASFKLRSFLEQFIYTTENTNFGN-RLV-----GIS	406
Db	321	NVSSIGSWYDKAPSGFVIESSVIRPP-HVFDYITGLTVYTSQRSISSARYIRHWAGHQS	379
Qy	407	NRDAPTVSNTITETLYGERTSGPTTKTIRPFESYKVSIVTDQSPVPSPIQP-----HFI	461
Db	380	YHRVSRGN--LQOMYGTNQLHSTSF-DFTNVDIYKTLSDKAVLLDIVPYGYVIFPG	436
Qy	462	INQIELYNGSSNN---TLKYSAGSLSNYQNTTFPFQPRKKDCNLVIDPQCGSPNPNYS	518
Db	437	MPEVEFFWQNLNTRKTLKY-----NPVSKDIIASTRDSLELPPETSDQPNYESYS	489
Qy	519	HILSHF-SLFTYSYVIGLQQLDGTGLVGNTHSSVDRYNAISDKIITMIPA-K-GNNLDT	576
Db	490	HRLCHITSIPATGNTTGL-----VPVFSWTHRSADLNNITYSDKITQIPAVKWCNDLPP	543
Qy	577	NSKVIIEGGHTGNLV-YLSQG-----RLEITCTPNSTQSYFRLRYATGNAGNT	627
Db	544	-VPVVGPGHGTGGDQLQYNRSTGVTGLFARYGLALE---KAGYRVRRLRYATDA----	595
Qy	628	LPNLSLTPIGVIGIPPORLANTFSGTNNYNNIQYGDGFYFQPPSTVTLPLNRNIPFIFNRA	687
Db	596	--DIVLHVNDQAQIQPKTMNPG-----EDLTSKTKVADAITTVNLATDSSVAVKHNLG	647
Qy	688	DVNSG----ILIIDKIEFPI 704	
Db	648	EDPNSTLSGIVYVDRIBFIPV 668	
RESULT 11			
US-11-108-389-74			
; Sequence 74, Application US/11108389			
; Publication No. US20050261188A1			
; GENERAL INFORMATION:			
; APPLICANT: Andre R. Abad			
; APPLICANT: Ronald D. Flannagan			
; APPLICANT: Rafael Herrmann			
; APPLICANT: Theodore W. Kahn			
; APPLICANT: Albert L. Lu			
; APPLICANT: Billy Fred McCutchen			
; APPLICANT: James K. Presnall			
; APPLICANT: James F.H. Wong			
; APPLICANT: Cao-Guo Yu			
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal			
; FILE REFERENCE: 35718/291049			
; CURRENT APPLICATION NUMBER: US/11/108,389			
; PRIOR FILING DATE: 2005-04-18			
; PRIOR APPLICATION NUMBER: 60/391,786			
; PRIOR FILING DATE: 2002-06-26			
; PRIOR APPLICATION NUMBER: 60/460,787			
; PRIOR FILING DATE: 2003-04-04			
; PRIOR APPLICATION NUMBER: 10/606,320			
; PRIOR FILING DATE: 2003-06-25			
; NUMBER OF SEQ ID NOS: 134			

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis (mutated)
US-11-108-389-74

Query Match      17.0%; Score 657.5; DB 7; Length 675;
Best Local Similarity 28.5%; Pred. No. 3.3e-43;
Matches 211; Conservative 120; Mismatches 293; Indels 117; Gaps 31;

QY 5 NDNVEYIIDSHTSPYFPNRSNDS--RYPYTNPNQPLONTNYKEWLNMCQGN--TQYGDN 62
Db 4 NNQNEYIIDATPS-----TSVSNDSNRYFPANFPTNALQNMKYDKYDKMSAGNASEYPGS 59
QY 63 FETFASADTAAVSAGTIVSGTLLAGIGLTSISGPIIGIIGAIISFGTLITVFWPAGEQ 122
Db 60 PEVLVSGQD--AKAAIDIVGKLSGLG-----VPFVGPIVSLYTQLIDILWPSGE- 108
QY 123 DKTWTQFIMGGEIFVDTPLETESIKQLKLTLEGFRQILQSINTALDDMRKRLQAPGL 182
Db 109 -KSQWEIFMEQVEELNQIAEYARNKALSEGLGNNGNYQLYLTALAEWEE--NPFERSRGF 166
QY 183 PPSALQQAALTUKIRFENVHNDPIREIPGFOLEYTKLLPIYAQAANFHLNLQOGAE 242
Db 167 RSGALRD-----VRNRFELDSLFTQYMPSFRTVNFPEVPFLTYVYAMAANLHLLKDAI 222
QY 243 LADEWNADIHPSQIEPNAGTSDDDYKLLKENIPKYSNYCANTYRTGLKNLDRDEPNMKWSI 302
Db 223 FGEWGW-----STTINNYDRQMKLTAEYSDHCVKWYETGLAKLGTSAKQWVD 273
QY 303 FNDYRYMTITVLDITISQFSLYDIKRYRDSIGGIEVKIGKINELTREIYTYTEINFDRLPQL 362
Db 274 YNQFREMTLAVLDVVALFPNYDTRY-----PMETKA---QLTREYVT-----DPLGAV 320
QY 363 RVQ-----PNLATMEYNLTRASPKLFSFLEQFIYFNTENTNFGN--RLV-----GIS 406
Db 321 NVSSIGSWYDKAPSGFVIESSVIRPP--HVFDYITGLTVYTSQSRISISSARYIRHWAGHQS 379
QY 407 NRDAPTYNTITELTGERTGSPTTKIRPESYKYSIVTDROSPVSPPIQ-----HFI 461
Db 380 YHRVSRGSN--LQOYGTQNQLHSTSTF--DFTNYDIYKTLSDKDAVLDDIVPGYTIFFG 436
QY 462 INQIELYLANGSSNN---TLKYSAGGSLSNQNTTFFQPRKDCNVLVIDPGCSNPNFNNYS 518
Db 437 MPEVEFFMVNQLNNTKTLKY-----NPVSKDIIASTRDSLELPPETSQDNYESYS 489
QY 519 HILSHF--SLFTYSYVIGLQQLDGTGVLGWTSSVDRYNAISDKIITMIPAIC--GNLNDT 576
Db 490 HRLCHITSIPATGNTTGL-----VPVFSWTHRSADLNNTIYSDKITQIPAVKWCNDLPP 543
QY 577 NSKVIEGPGHTGNLV--YLQSQ-----RLEITCETPNSTQSFIRLRYATNGAGNT 627
Db 544 -VPVKGPGHTGGDLLQYNRSTGVSGLTFLARYGLALE---KAGKYRVLRYATDA---- 595
QY 628 LPNISLTI PGVIGIPQRLNNTSGTNNYNLQYDGFYQFPSTVTPLPLNRNIPFIFNRA 687
Db 596 --DIVLHVDAQIQMKTWNP-----EDLTSKTFKVAADAITVNLATSSVAVKHNLG 647
QY 688 DVSNS----ILIIDKIEFPI 704
Db 648 EDPNSTLSGIVVYDRIEFIPV 668

RESULT 12
US-11-108-389-80
; Sequence 80, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
```

```
Db      596  --DIVLHVNDQAQIOMPKTNPG-----EDLTSKTFKVADAITVNLATDSSVAVKHNLG 647
Qy      688  DVSN-----ILIIDKIEFPI 704
Db      648  EDPNSTLSGIVVVDRIEFIPV 668

RESULT 13
US-11-058-727-82
; Sequence 82, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis (mutated)
US-11-058-727-82

Query Match      17.0%; Score 657; DB 7; Length 674;
Best Local Similarity 28.5%; Pred. No. 3.6e-43;
Matches 212; Conservative 118; Mismatches 291; Indels 122; Gaps 31;

Qy      5      5 NDNNEYIIDSHTSPYFPNRSNDS-RYPYTNPNQPLQNTNYKEWLNMCQGN-TOYGDN 62
Db      4      4 NNQNEYIIDATPS-----TSVSDNSNRYPFANEPTNALQNDYKYDKLXMSAGNASEYPGS 59
Qy      63      63 FETPASADTTAAVSAGTIVSGTLGAGIGLTSISGPIGIIIGAIISFGTLITVFWPAGEQ 122
Db      60      60 PEVLVSGQD--AAKAAIDIVGKLISGLG-----VPFVGPIVSLYTLQILDLWPSGE- 108
Qy      123      123 DKTWVTOFKMGEIFVDTPLTSTESIKQLKLTLEGFRQLQSINTALDDWRKLRQAPGL 182
Db      109      109 -KSQWEIFMEQVEELINOKIAEYARNKALSLEGLGNQYQLYLTALBEEWEE-----N 159
Qy      183      183 PPSSAL--QQAALTKEIFENVHNDREIPGFOLEYTKTLPIYAQAANFHLNLQOG 240
Db      160      160 PNGSRFRQALRDVRNRFLDLSLFTQYMPFSRVTNFEVFLTVYAAANLHLULLKDA 219
Qy      241      241 AELADEWNADIHPQIBFNAGTSDYYKLLKENIPKYSNYCANTYRTGLKMLRDEPNMKW 300
Db      220      220 SIFGEENGW-----SITTIINYYDRQMKLTAEYSDHCVKWYETGLAKLKGTSKQW 270
Qy      301      301 SIFNDYRYMTITVLDTSOFSYDIKRYRDSIGGI EVKGIKBLTREIYTFTEINFRRLP 360
Db      271      271 VDNQPRREMTLAVLDVVALFPNDYTRTY-----PMETKA---QLTREYVT-----DPLG 317
Qy      361      361 QLRVQ-----PNLATMEYNLTRASFKLFSLEQEIFYTTENTNFCN-PLV-----G 404
Db      318      318 AVNVSSIGSWYDKAPSGVIESSVIRPP-HVFDYITGLTVYTQSRSSIRYIRHWAGHQ 376
```

```
Qy      405      405 ISNRDAPYTSNTITETLYGERTGSPTTKIRPEFSKVSIIVTDROSPVPPIOP-----H 459
Db      377      377 ISYHRVSRGSN--LQOMYGTNQNHLSTFP-DFTNYDIYKTLUSKQAVLDIDIVPGVYIYF 433
Qy      460      460 FIINQIELYLMGSSNN---TLKYSAGGSLSNQNTTFFQPRKKOCNCLVIDPGCSFNFN 516
Db      434      434 FGMPEVEFFMVNQLNNTRKTLKY-----NPVSKDIIASTRDSLELEPETSDDQNYES 486
Qy      517      517 YSHLSHP-SLPTYSYVIGLQILDTGVLGTHSSVDRYNAISDKIITMIPAUK-GNNL 574
Db      487      487 YSHRLCHITSIPATGNTTGL-----VPVFSWTHRSADLNNTIYSDKITQIPAVKCDNL 540
Qy      575      575 DTNSKVIKSGHGTGNLV-YLQSQG-----RLEITCETPNSQTSYFIRLRYATNGAG 625
Db      541      541 PF-VPVVGPGHGTGDLLOYNRSTGVTGLFLARYGLALE---KAGKYRVRURYATDA-- 594
Qy      626      626 NTLNPNISLITPGVIGIPQRLNNTFSGTNNYNLQYDFGFGYFQFPSTVTLPLNRNIPFIN 685
Db      595      595 ----DIVLHVNDQAQIOMPKTNPG-----EDLTSKTFKVADAITVNLATDSSVAVKH 644
Qy      686      686 RADVSNS----ILIIDKIEFPI 704
Db      645      645 LGEDPNSTLSGIVVVDRIEFIPV 667

RESULT 14
US-11-108-389-82
; Sequence 82, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis (mutated)
US-11-108-389-82

Query Match      17.0%; Score 657; DB 7; Length 674;
Best Local Similarity 28.5%; Pred. No. 3.6e-43;
Matches 212; Conservative 118; Mismatches 291; Indels 122; Gaps 31;

Qy      5      5 NDNNEYIIDSHTSPYFPNRSNDS-RYPYTNPNQPLQNTNYKEWLNMCQGN-TOYGDN 62
Db      4      4 NNQNEYIIDATPS-----TSVSDNSNRYPFANEPTNALQNDYKYDKLXMSAGNASEYPGS 59
Qy      63      63 FETPASADTTAAVSAGTIVSGTLGAGIGLTSISGPIGIIIGAIISFGTLITVFWPAGEQ 122
Db      60      60 PEVLVSGQD--AAKAAIDIVGKLISGLG-----VPFVGPIVSLYTLQILDLWPSGE- 108
Qy      123      123 DKTWVTOFKMGEIFVDTPLTSTESIKQLKLTLEGFRQLQSINTALDDWRKLRQAPGL 182
Db      109      109 -KSQWEIFMEQVEELINOKIAEYARNKALSLEGLGNQYQLYLTALBEEWEE-----N 159
```

```
QY 183 PPSSAL--QQAALTAKIRPENVNDPIREIPGQLEYTKYKTLPIYQAANFHLNLLQOG 240
Db 160 PNGSRFRSQALDRVNRFEILDSLTQYMPSPRVTFNFEVPLFTVYAMAANLHLLKDA 219
QY 241 AELADEWNAIDHPSQIEPNAGTSDYYKLLKENIPKYSNYCANTYRTGLKNLKDPEPMKW 300
Db 220 SIFGEWGW-----STTTNNYDROMKLTAEYSDHCVKWYETGLAKLKGTSKQW 270
QY 301 SIPNDYRRYMTITVLDITISQFSLYDIKRYRDSIGGIEVGKIKNELTREIYTTREINFDRLP 360
Db 271 VDYNQPERENTLAVLDVVALFPNYDRTY-----PMETKA---QLTREYVT---DPLG 317
QY 361 QLRVQ-----PNLATMEYNLTASFKLFSLEQFIYTYTNTNFGN-RLV-----G 404
Db 318 AVNVSSIGSWYDKAPSGFVIESSVIRPP-HVFDYITGLTVYTSRSISSARYIRHWAGHQ 376
QY 405 ISNRDAPTYNTITETLYGERTGSPTKTIRPESYKVSIVDRQSPVPSIOP-----H 459
Db 377 ISYHRVSRGSN--LQOMYGTNQLHSTSTF-DFTNYDIYKTLSDKDAVLLDIVYPGYTYIF 433
QY 460 FIINQIELYNGSSNN---TLKYSAGSLSNYQNTTFFQPRKDCNLDVIDPGCSNPNFN 516
Db 434 FGMPEVEFPWVNLNTRKTLKY-----NPVSKDIIASTRSELELPETSDQPNYES 486
QY 517 YSHLSHP-SLFTYSYVIGLQILDTGVLGWTSHSSVDRYNAISDKIITMIPAUK-GNNL 574
Db 487 YSHRLCHITSIPATGNTTGL-----VPVFSWTHRSADLNNNTIYSDKITQIPAVKCDNL 540
QY 575 DTNSKVIIEPGHGTGNLV-YLOSQ-----RLBITCTPNSTQSYFIRLRYATNGAG 635
Db 541 PF-VPVVKPGHGTGGLLQYNRSTGVTGLFLARYGLALE---KAGKYRVLRYATDA-- 594
QY 626 NTLPNLSLTIPGVIPIPPORLANTFSGTNVNNLOYGDFGVFPSTVTLPLNRPNIPIFN 685
Db 595 ----DIVLHVNDQAIOQPKTMNPG-----EDLTSKTFKVADAITTVNLATDSSVAVKH 644
QY 686 RADVNS-----ILIIDKIEFPI 704
Db 645 LGEDPNSTLSGIVYVDRIEFIV 667
```

RESULT 15

```
US-11-058-727-42
; Sequence 42, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE REFERENCE: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis (mutated)
```

US-11-058-727-42

```
Query Match 17.0%; Score 656.5; DB 7; Length 675;
Best Local Similarity 28.5%; Pred. No. 3.9e-43;
Matches 211; Conservative 120; Mismatches 293; Indels 117; Gaps 31;

QY 5 NDNNEVEIIDSHTSPYFPNRSNDS-RYPYTNNNQPLONTNYKEWLNMCQGN-TOYGDN 62
Db 4 NNQNEYBIIDATPS---TSVSNDSNRYFPANEPNTALQNNMDYKDYKMSAGNASEYFCS 59
QY 63 FETPASADTTAAVSAGTIVSGTILAGIGLTSISGPIIGIIGAIISFOTLITVFPAGEQ 122
Db 60 PEVLVSQD--AKAAADIVGKLLSLG-----VPFVGPIVSLYTQLDILWPSGE- 108
QY 123 DKTWTQFIKMGEIFVDPTTESIKQLKQTLQLEGFRQILOQSYNTALDDWRKLRLQAPGL 182
Db 109 -KSQWEIFMQVEELINQIAEYARNKALSELEGNGNYQLYLTALBEEW- NPPRSRGF 166
QY 183 PPSALQQAALTUKIRPENVNDPIREIPGQLEYTKYKTLPIYQAANFHLNLLQOGAE 242
Db 167 RSRGALRD---VRNRFEILDSLTQYMPSPRVTFNFEVPLFTVYAMAANLHLLKDAI 222
QY 243 LADEWNAIDHPSQIEPNAGTSDYYKLLKENIPKYSNYCANTYRTGLKNLKDPEPMKWI 302
Db 223 FGBEWGW-----STTTNNYDROMKLTAEYSDHCVKWYETGLAKLKGTSKQWVD 273
QY 303 FNDYRRYMTITVLDITISQFSLYDIKRYRDSIGGIEVGKIKNELTREIYTTREINFDRLP 362
Db 274 YNQFRREMTLAVLDVVALFPNYDRTY-----PMETKA---QLTREYVT---DPLGAV 320
QY 363 RVQ-----PNLATMEYNLTASFKLFSLEQFIYTYTNTNFGN-RLV-----GIS 406
Db 321 NVSSIGSWYDKAPSGFVIESSVIRPP-HVFDYITGLTVYTSRSISSARYIRHWAGHQIS 379
QY 407 NRDAPTYNTITETLYGERTGSPTKTIRPESYKVSIVDRQSPVPSIOP-----HFI 461
Db 380 YHRVSRGSN--LQOMYGTNQLHSTSTF-DFTNYDIYKTLSDKDAVLLDIVYPGYTYIF 436
QY 462 INQIELYNGSSNN---TLKYSAGSLSNYQNTTFFQPRKDCNLDVIDPGCSNPNFNYS 518
Db 437 MPEVEFPWVNLNTRKTLKY-----NPVSKDIIASTRSELELPETSDQPNYESYS 489
QY 519 HILSHP-SLFTYSYVIGLQILDTGVLGWTSHSSVDRYNAISDKIITMIPAUK-GNNLDT 576
Db 490 HRLCHITSIPATGNTTGL-----VPVFSWTHRSADLNNNTIYSDKITQIPAVKCDNL 543
QY 577 NSKVIIEPGHGTGNLV-YLOSQ-----RLBITCTPNSTQSYFIRLRYATNGAGNT 627
Db 544 -VPVVKPGHGTGGLLQYNRSTGVTGLFLARYGLALE---KAGKYRVLRYATDA---- 595
QY 628 LPNLSLTIPGVIPIPPORLANTFSGTNVNNLOYGDFGVFPSTVTLPLNRPNIPIFNRA 687
Db 596 --DIVLHVNDQAIOQPKTMNPG-----EDLTSKTFKVADAITTVNLATDSSVAVKH 647
QY 688 DVNS-----ILIIDKIEFPI 704
Db 648 EDPNSTLSGIVYVDRIEFIV 668
```

Search completed: December 4, 2005, 13:04:35

Job time : 14 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 11, 2005, 22:52:13 ; Search time 9651 Seconds

(without alignments)
4329.079 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869

Sequence: 1 MNQNDNNEVEIIDSHSPV.....KLETIQKINTFTNHTKTL 735

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10783417/runat_01122005_141956_24331/app_query.fasta_1.903
-DB=GenEmbl -OFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPECI=0 -LOOPEXT=0
-UNITS=bites -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NOR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10783417 @CGN 1 1 6416 @runat_01122005_141956_24331 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-FGV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sts.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3869	100.0	2208	6	CQ868305 Sequence
2	3314	85.7	2235	6	CQ868307 Sequence
3	3104.5	80.2	2085	6	CQ868309 Sequence

	4	1061.5	27.4	3543	1	BTTOXD1
5	1061.5	27.4	4451	6	I08083	Y00423 Bacillus th
6	1061.5	27.4	4451	6	I09103	I08083 Sequence 1
7	1060.5	27.4	4253	6	BACISRH4	D00248 Bacillus th
8	1060.5	27.4	4934	6	E01676	E01676 DNA sequenc
9	1060.5	27.4	127923	1	BTBPTOXIS	AL731825 Bacillus
10	897.5	23.2	3753	1	BACMSQB	ML2662 B.thuringie
11	897.5	23.2	127923	1	BTBPTOXIS	AL731825 Bacillus
12	880	22.7	3756	6	E00614	E00614 DNA encodin
13	862.5	22.3	4959	1	AF285775	AF285775 Bacillus
14	850	22.0	6009	1	BTB251977	AJ251977 Bacillus
15	849	21.9	4896	1	AF132928	AF132928 Bacillus
16	808.5	20.9	3752	1	AB125059	AB125059 Bacillus
17	784	20.3	3644	1	BTB251978	AJ251978 Bacillus
18	775.5	20.0	3507	1	BTU04365	U04365 Bacillus th
19	775.5	20.0	3507	6	I25972	I25972 Sequence 3
20	759	19.6	3684	1	BTTOXD2	X07423 Bacillus th
21	759	19.6	3684	6	I08884	I08884 Sequence 2
22	757.5	19.6	3535	6	I06096	I06096 Sequence 2
23	754	19.5	3684	1	BTITOX	X07082 Bacillus th
24	752.5	19.4	4186	1	BACISRH3	D00247 Bacillus th
25	752.5	19.4	4186	6	E01905	E01905 genomic DNA
26	746	19.3	3642	1	AB193814	AB193814 Bacillus
27	709	18.3	3668	1	D88381	D88381 Bacillus th
28	705	18.2	4391	1	BTBEG65	Y07603 B.thuringie
29	705	18.2	6698	1	AB116651	AB116651 Bacillus
30	703	18.2	3633	6	CS130948	CS130948 Sequence
31	703	18.2	3633	6	AX543926	AX543926 Sequence
32	703	18.2	6613	6	CS130963	CS130963 Sequence
33	703	18.2	6613	6	AX543951	AX543951 Sequence
34	696.5	18.0	4056	1	BACCRID2	M20242 B.thuringie
35	696	18.0	3536	6	E01029	E01029 DNA sequenc
36	689.5	17.8	4188	6	BD177494	BD177494 Polypepti
37	689.5	17.8	4188	6	AX670919	AX670919 Sequence
38	687	17.8	3931	1	AB161456	AB161456 Bacillus
39	678.5	17.5	2003	6	AX543930	AX543930 Sequence
40	678.5	17.5	2022	6	CS130958	CS130958 Sequence
41	678.5	17.5	2022	6	AX543940	AX543940 Sequence
42	677	17.5	8572	1	AB116649	AB116649 Bacillus
43	677	17.5	8605	1	AB116650	AB116650 Bacillus
44	676	17.5	1953	6	I25973	I25973 Sequence 5
45	676	17.5	1977	1	BTCADE	X17123 Bacillus th

ALIGNMENTS

RESULT 1	CQ868305	Sequence 6 from Patent WO2004074462.	2208 bp	DNA	linear	PAT 13-SEP-2004
LOCUS	CQ868305	Sequence 6 from Patent WO2004074462.	2208 bp	DNA	linear	PAT 13-SEP-2004
DEFINITION	CQ868305	Sequence 6 from Patent WO2004074462.	2208 bp	DNA	linear	PAT 13-SEP-2004
ACCESSION	CQ868305	Sequence 6 from Patent WO2004074462.	2208 bp	DNA	linear	PAT 13-SEP-2004
VERSION	CQ868305.1	GI:51998351	2208 bp	DNA	linear	PAT 13-SEP-2004
KEYWORDS	CQ868305.1	GI:51998351	2208 bp	DNA	linear	PAT 13-SEP-2004
SOURCE	CQ868305.1	GI:51998351	2208 bp	DNA	linear	PAT 13-SEP-2004
ORGANISM	CQ868305.1	GI:51998351	2208 bp	DNA	linear	PAT 13-SEP-2004
REFERENCE	CQ868305.1	GI:51998351	2208 bp	DNA	linear	PAT 13-SEP-2004
AUTHORS	CQ868305.1	GI:51998351	2208 bp	DNA	linear	PAT 13-SEP-2004
TITLE	CQ868305.1	GI:51998351	2208 bp	DNA	linear	PAT 13-SEP-2004
JOURNAL	CQ868305.1	GI:51998351	2208 bp	DNA	linear	PAT 13-SEP-2004
FEATURES	CQ868305.1	GI:51998351	2208 bp	DNA	linear	PAT 13-SEP-2004
source	CQ868305.1	GI:51998351	2208 bp	DNA	linear	PAT 13-SEP-2004
CDS	CQ868305.1	GI:51998351	2208 bp	DNA	linear	PAT 13-SEP-2004

/translation="MNQNDNNEYEIIIDSHTSYPFNENSDSRYPYNNPNQPLQNT
NYKWLNMCGQNTYQNDQNFETFSADTIIVSGTLLAGIGLISISGPIGII
GAILISFGTLITVFWAGQDQTVMTQIFKMGSEIFVDYPLTESIKOLKQTLGEGFQI
LOSNTALDDWRKRLQLOAPGLPSSALQQAALTKIRFENVHNDFTIREIPGQLETY
KTLILLPIYAAQANFLNLQOQGAELADENWADTHPSQIBNAGTSDDYKLLKENIPK
YSNYCANVTGTGLKNLRDNPENMKWSIFNDYRRVMTITVLDTISQFSLYDKRVRDSIG
GLEVGKIKNELREIYTHEINFDRLPQLRVQPNLATMEYNLTRASPKLPSFLQDFIF
TENINFGRLVLGINSDDAPYSNTITETLYGERTGPTTKTIRPFESYKVSIVTDQOS
PPVSPLOPHFIINQIELYNGSNNTLYKSAGSLNQNTTFFQPRKDCNLDVDP
GSGPNFNNTSHILSLFTYSYVIGLOQLIDTGLVWTHSSVDRYNAISDKIITMI
PAIKNNLDTNKSIVGPGHGTGNLVLOSQGRLEITCETPNSTQSVYFRLRYATNGA
GNTLPNLSITPGVIGIPORLNTPSGTNNYNLOYDGVGFQFPPSTVTLPLNRNIPF
INRADVSNLSILLIDKIEPIPTSSMHQREKQKLETIQTKINTFFFNHFKTL"

ORIGIN

Alignment Scores:
Pred. No.: 5.4e-283 Length: 2208
Score: 3869.00 Matches: 735
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-783-417-2 (1-735) x CQ868305 (1-2208)

QY 1 MetAsnGlnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
DB 1 ATGAATCAAAATAACGATAATAACGAATATGAAATATTATTCATCGGCATACCTCACCTTAT 60
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
DB 61 TTTCGGAACAGAAACAGTAGTATGATTCCTAGATACCTTTACACAAATAATCAAAATCAACCA 120
QY 41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
DB 121 TTACAAACACAAATTACAAAGAGTGCTCAATATGTGTCAAGGGAATACCAATATGGT 180
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80
DB 181 GATAATTCGAGACATTTGCTAGTGTGATGACAAATTTGCTGAGTGTAGTCAGTACTATT 240
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
DB 241 GTATCCGGTACTCTGTTAGCCGTATAGGTGGGCTCACTTCATATCCGACCGATAGGA 300
QY 101 IleIleGlyValIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGly 120
DB 301 ATAAATAGGTGCTATAATAATATCTTTTGGTACCCTTAATCACTGTCTTTTGGCCCGGGGA 360
QY 121 GluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
DB 361 GAACAAGACAAACAGTATGGACACAAATTTATTAATAATGGAGAAATTTTGTGTGATACA 420
QY 141 ProLeuThrGluSerLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
DB 421 CCGTTAACAGAAAGCATAAACAGCTAAAGTTACAAACTTTAGAAAGATTAGACAAATA 480
QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTyrPArgLysLeuLysArgLeuGlnAlaPro 180
DB 481 TTACAAAGCTATTAATACAGCATTTAGATTTGGAGAAATTTAAAGAGACTTACAAGCTCCT 540
QY 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGlu 200
DB 541 GGATTAACCATCATCAGCATTTACAAACAGCTGCGCTTGACTCTTAATAATACGATTTGAG 600
QY 201 AsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThr 220
DB 601 AATGTTTCACAATGATTTTATTCAGAAATACCTGGTTTCCAACTTGAAACTTATAAAACG 660
QY 221 LeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGly 240
DB 661 CTATTACTACCTATTATTATGCGCAAGCTGCTAAATTTTCAATTTAAATTTATTATACCAACAGGT 720

QY 241 AlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAla 260
DB 721 GCTCAATTTGCTGATGAATGAATGCAATATACATCTTCCACAAATTTGAACCTAATGCT 780
QY 261 GlyThrSerAspAspTyrTyrLysLeuLysGluAsnIleProLysTyrSerAsnTyr 280
DB 781 GGAACATCAGATGACTATTATAAACTTTTAAAGAAATATACCTTAATATATAGTAACCTAT 840
QY 281 CysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTyr 300
DB 841 TGTGCAAAATACCTATAGAACAGGACTAAAAATCTTAGAGACGACCAAAATATGAAATGG 900
QY 301 SerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGln 320
DB 901 AGTATATTTAATGACTATCGAAGATATATGACCATTATCTGTATTAGATACCATCTCTCAA 960
QY 321 PheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluValLysGly 340
DB 961 TTTTCTTTTATGATATAAAAGATATAGAGATTCAATAGGAGGAATAGAAAGTAAAGGC 1020
QY 341 IleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuPro 360
DB 1021 ATTAAGAAATCACTCACAGAGAAATTTATACAACTGAATATAATTTTGATCGTCTTCT 1080
QY 361 GlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPhe 380
DB 1081 CAACCTTAGAGTTCAACCCCAATCTAGCTACGATGGAATATAATTTAAACACGTCGAAGTTTT 1140
QY 381 LysLeuPheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsn 400
DB 1141 AAATATATTTTCATTTTATAGAACAAATTTATTTTATACAGAAATACAAATTTTCGGGAAT 1200
QY 401 ArgLeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
DB 1201 CGTTTAGTGTGATTTCTTAATCGTAGTGACCTACTTATAGCAATATCTATTAACCTGAACT 1260
QY 421 LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440
DB 1261 TTATATGGAGAAAGAACAGGTTCACCCACACAAACAAATAAGACCATTTTGAATCTTAT 1320
QY 441 LysValSerIleValThrAspArgGlnSerProValSerProIleGlnProHisPhe 460
DB 1321 AAAGTTTCAATTTGATTAATGATAGACAAATCACCTCTCTCTTCCCTATTTCAACCACTTT 1380
QY 461 IleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsnThrLeuLysTyrSer 480
DB 1381 ATAAATTAATCAAAATTTGAACTTTATTTAATGGCTCATCTAACACACACACTCAATATTTCA 1440
QY 481 AlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArgLysLys 500
DB 1441 GCAGGAGGGTCTTTATCTAATTTACAAACACAACTTTTTCATTTTCTTAGAAAAA 1500
QY 501 AspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIle 520
DB 1501 GACTGCAATCTAGTATTGATTCAGGTGTTCACCAAACTTTAATACTATAGTCATATT 1560
QY 521 LeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyGlnLeuGlnIleLeu 540
DB 1561 TTATCCCAATTTTCAATTTTACTTATCTCTGTGATTTGATTTACAGCTACAAATATTA 1620
QY 541 AspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerAsp 560
DB 1621 GATACAGGTGTATTAGATGGACACACAGTAGTGTGTGATAGATATAATGCAATATCAGAT 1680
QY 561 LysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysVal 580
DB 1681 AAAATAATTAACAATGATCCAGCAATCAAAGGTAAACAATCTTGATACAAACTCTAAGGTA 1740
QY 581 IleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeu 600
DB 1741 ATTTGAAGGAGCTCGTTCATACAGGAGGAACTTGGTTTTTATTTACAAAGTCAAGGGCGTTTA 1800
QY 601 GluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAla 620


```

|||||
1801 GAAATTACATGTGAACTCCTAAATTCACACAATCTTATTTCATAGACTTCGATATGCT 1860
QY 621 ThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGly 640
Db 1861 ACAAAATGGTCTGGAATACTCTCTCTTAATATATCTCTTCAATATACAGAGTAAATAGGA 1920
QY 641 IleProGlnArgLeuAsnAsnThrPheSerGlyThrAsnThrLeuAsnLeuGlnThr 660
Db 1921 ATACCACTCAGAGACTCAACAACACTTTCTCGTACAAATATATAATTTTCAATAC 1980
QY 661 GlyAspPheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIle 680
Db 1981 GGAGATTTTGGGTATTTCCAAATTTCCAAATGACAGTACAGTAACTTACCTTTAAATCGAACA 2040
QY 681 ProPheIlePheAsnArgAlaAspValSerAsnSerIleLeuIleAspLysIleGlu 700
Db 2041 CCATTTATATATTAATCTGTGAGATGTATCAAAATTCAAATTTTAAATCAATGATAAAATTGAA 2100
QY 701 PheIleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIle 720
Db 2101 TTTATACCAATTTACTCTCTATGCACCAAAATAGAGAAAACAAAATTAGAAACTATC 2160
QY 721 GlnThrLysIleAsnThrPhePheThrAsnHisThrLysThrLeu 735
Db 2161 CAACAATAATAATACATTTTTCACAAATCATACAAAACACTT 2205

RESULT 2
CQ868307 2235 bp DNA linear PAT 13-SEP-2004
LOCUS Sequence 8 from Patent WO2004074462.
CQ868307 8 from Patent WO2004074462.
CQ868307.1 GI:51998353
VERSION 1
KEYWORDS Bacillus thuringiensis
SOURCE Bacillus thuringiensis
ORGANISM Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.

REFERENCE
1. Carozzi, N., Hargiss, T., Koziel, M.G., Duck, N.B. and Carr, B.
Delta-endotoxin genes and methods for their use
Patent: WO 2004074462-A 8 02-SEP-2004;
Athenix Corporation (US)

FEATURES
source
1. .2235
/organism="Bacillus thuringiensis"
/mol_type="unassigned DNA"
/db_xref="taxon:1428"
CDS
1. .2235
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAH33949.1"
/db_xref="GI:51998354"
/translations="MKQNNNEVEIIDSKNLSPSNRNIDHSRIPTYNNQPLQNTN
YKEWLNMCQNTQYGNFENFASADITAAVSAGTIVSGTLGAGIGLTSISGPIIG
AIIISFGLITVFWPGEQDKWTQFIKMEIFVDTPLTESIKQLKLOTLEGFQIL
QSYNTALDDWRKLQAPGLPPSSALQQAALTKIRFENVHNDFTREIPGFOLETYK
TLALPIYAQAANFHLNLOQAEALDEWNAIDHPSQIEPNAGTSDDYKLLKENIPKY
SNYCANTYREGLKLNENPMRWSIFNDYRYMTITVLDITIAQFSFVDIKRYKDSIGR
IGGKTELREIYITEINFORLTYLEIQPNLAIMENYLNTRSGRLSFLDELIFYTKN
ETVGNLAVGANRSTYATGTTEIYIGERTGPTTKLIPFESYKVSIVTDQVPT
SPPNYFTNQILELYNNSPKLTYSAGNLNDKKTDFQFPVKKQCKPNIINPC
LPNSYNSHTLSQFSPNYSYKIGLAINILYTGALGWTSSVNRNNAISDKIITWIPA
IKGNSLDTSNKVTEGEGHTGGLNLYLQSQGLEITCETPNSTQSYIRLYRYNAGN
TLNLSITLPGVIGIPQRLNLYFGNINNLQIDGFYQFSEFTVLPNRPNIPFI
NRADVNSILIIDIKIEFIPITTSVRQNRQKQLETTQTKINTFTTNHNTNLINATN
YDID"

ORIGIN
Alignment Scores: 5.39e-241 Length: 2235
Pred. No.: 3314.00 Matches: 638
Score:

```

```

Percent Similarity: 91.28% Conservative: 32
Best Local Similarity: 86.92% Mismatches: 60
Query Match: 85.66% Indels: 4
DB: 6 Gaps: 3
US-10-783-417-2 (1-735) x CQ868307 (1-2235)

QY 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 1 GTGAATCAAAATAAT--ATAATGAATATGAGATTATCGATTCAAGAAATTTATCTTAT 57
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 58 CTTTCTAACAGAAATATTGATCATTTCTAGATACCTTTACACAAATAATCCAAATCAACA 117
QY 41 LeuGlnAsnThrAsnTyrLysGluThrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
Db 118 TTACAAACACAAATTTACAAAGAGTGGCTCAATATGTCTCAAGGGAATACACAATATG 177
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaIaValSerAlaGlyThrIle 80
Db 178 GATAAATTTTCAGACATTTGCTAGTGTGATACAAATGCTGAGTTAGTGCAAGGTACTATT 237
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
Db 238 GTATCCGGTACTCTGTTAGCCGATATAGTGGGCTCACTTCTATATCCGACCGATAGA 297
QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120
Db 298 ATAATAGTGTCTATAATAATATCTTTTGGTACCCCTAATCACTGCTCTTTTGGCCGCG 357
QY 121 GluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db 358 GAACAAGACAAAACAGTATGACACAAATTTATTAATAATGGGAGAAATTTTGTGTGATA 417
QY 141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
Db 418 CCGTTAACAGAAAGCATAAACACAGCTAAAGTTTACAACTTTTAGAAGGATTTAGACAA 477
QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaPro 180
Db 478 TTACAAAGCTATATAATACAGCATTTAGATGATGGAGAAAATTAATAAGACTACAGCT 537
QY 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGlu 200
Db 538 GGATTACCACTCATCAGCATTTACAAACAGCTGCTTGAATCTTTAAATAATACGATTTG 597
QY 201 AsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThr 220
Db 598 AATGTTCCAAATGATTTTATTCGAGAAATACCTGGTTTCCAACTTGAACACTTATAA 657
QY 221 LeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGly 240
Db 658 CTATTACTACTATTATTCGCAAGCTGCTAATTTTCATTAAATTTATTACACAAAGT 717
QY 241 AlaGluLeuAlaAspGluThrProAsnAlaAspIleHisProSerGlnIleGluProAsnAla 260
Db 718 GCTGAATGGCTGATGAATGGAATGCAGATATACATCTCTTCAAAATTTGAACCTAATG 777
QY 261 GlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyr 280
Db 778 GGAACATCAGATGATATTATAAACTTTTAAAGAAAATATACCTTAAATATATAGTAAC 837
QY 281 CysAlaAsnThrTyrArgThrGlyLysLeuLysAsnLeuArgAspGluProAsnMetLysTrp 300
Db 838 TGTGCAAAATACCTATAGAGAGGACTAAATAAATTCGAAACGAACTTAATATAGATGG 897
QY 301 SerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGln 320
Db 898 AGTATATTTTAAATGATTATCGAAGATATATGACTATTACTGTATAGATACATCGCTCA 957
QY 321 PheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyIleGluValLysGly 340
|||||

```

Db 958 TTTTCTTTTATGATATAAAGAGATACAAAGATTCAATAGGA-----AGAATAGGTGGC 1011
Qy 341 IleLysAsnGluLeuThrArgGluIleTyrThrGluIleAsnPheAspArgLeuPro 360
Db 1012 ATTAACCTGAACCTTACAGAGAAATTTATACAACTGAAATAATTTTGACCGCTTACT 1071
Qy 361 GlnLeuArgValGlnProLeuLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPhe 380
Db 1072 TACCTTGAAATCAACCCAACTCGCTATATATGGAATATAATTTAAACACGTCAGCGCTT 1131
Qy 381 LysLeuPheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsn 400
Db 1132 AGATATTTTCAATTTTATGGAACCTATATTTATACAAAATGAAACGTCAGCGAAT 1191
Qy 401 ArgLeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
Db 1192 CGTTTAGTTGGTATTCGGAATCGTAATAGATCTACTTATCTACGACAGGAATGAAAT 1251
Qy 421 LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440
Db 1252 ATATATGGAGAAAGAACAGCTCCACCCACACAAACCTTTAATACCAATTTGAATCCTAT 1311
Qy 441 LysValSerIleValThrAspArgGlnSerProProValSerProIle----GlnProHis 459
Db 1312 AAAGTTTCAATTTGAATCTGATAGACAAAGTAACCTCTACTTCCCGCTTTTCCCTAACATATAC 1371
Qy 460 PheIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsnThrLeuLysTyr 479
Db 1372 TTTACAATTAATCAAAATGAACTTTATTTAAATAATTCACCTAGTATAATAATTAACATAT 1431
Qy 480 SerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPheGlnPheProArgLys 499
Db 1432 TCAGCTGGGGGAAATTAATCAATGATAAAACCAACTGATTTCAATTTCTGTGATAA 1491
Qy 500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHis 519
Db 1492 AAAGACTGTAACCAATTAATTAATCAAAATGTTTACCAAGCTATAATAGTTATAGTCAT 1551
Qy 520 IleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIle 539
Db 1552 ATTTTATCCAGTTTCTTTATTAATTAATATTCCTATAAAATGGATAGCGCTAAATATA 1611
Qy 540 LeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSer 559
Db 1612 TTATATACAGTGCATAGATGACACACAGTAGTGTAAATAGATAATATGCAATATCA 1671
Qy 560 AspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLys 579
Db 1672 GATAAAATRAATACAATGATCCAGCAATCAAGGTAACAGTCTTGATACAAACTCTAAG 1731
Qy 580 ValIleGluGlyProGlyHisThrGlyAsnLeuValTyrLeuGlnSerGlnGlyArg 599
Db 1732 GTAATTAAGAGGACCTGGTCAACAGGAGAACTTGTTTATTATACAAATCAAGGCGT 1791
Qy 600 LeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyr 619
Db 1792 TTAGAGATTACATGTAGAACTCTTAATTCACACAATCTTATACATAGACTTCGATAC 1851
Qy 620 AlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIle 639
Db 1852 GCTACAAATGGTCTGGAATACTCTTCCTAATATATCTCTTACAAATACCAGGAGTAATA 1911
Qy 640 GlyIleProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGln 659
Db 1912 GGAATACCACTCAACGACTCAACACACTTTTTCGTGTACAAATTAATAATTTTACAA 1971
Qy 660 TyrGlyAspPheGlyTyrPheGlnPheProSerThrThrValThrLeuProLeuAsnArgAsn 679
Db 1972 TACGGAGATTTTGGGTATTTCCAAATTTCCAAATGACAGTAAATACCTTTAAATCGAATC 2031
Qy 680 IleProPheIlePheAsnArgAlaAspValSerAsnSerIleLeuIleAspLysIle 699
Db 2032 ATACCATTTATATTTTAAATCGTCAGATGTATCAAAATTTCAATTTTAAATCAATTTGATAAAAT 2091

Qy 700 GluPheIleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThr 719
Db 2092 GAATTTATACCAATTAATCTCTCTGTAGCCAAATAGAGAAAACAAATAATGAAACT 2151
Qy 720 IleGlnThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db 2152 ATCAAAACAAAATAAATACATTTTTCACAAATCATACAAA 2193
RESULT 3
CQ868309
LOCUS CQ868309 2085 bp DNA linear PAT 13-SEP-2004
DEFINITION Sequence 10 from Patent WO2004074462.
ACCESSION CQ868309
VERSION CQ868309.1 GI:51998355
KEYWORDS Bacillus thuringiensis
SOURCE Bacillus thuringiensis
ORGANISM Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE 1
AUTHORS Carozzi, N., Hargiss, T., Koziel, M.G., Duck, N.B. and Carr, B.
TITLE Delta-endotoxin genes and methods for their use
JOURNAL Patent: WO 2004074462-A 10 02-SEP-2004;
Athenix Corporation (US)
FEATURES
Location/Qualifiers
source 1..2085
/organism="Bacillus thuringiensis"
/mol_type="unassigned DNA"
/db_xref="taxon:1428"
CDS 1..2085
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAH33950.1"
/db_xref="GI:51998356"
/translation="MCQGNQYQDNFETASADTAAVSAAGTIVSGTLLAGIGLTSI
SGPIGIIIGALIIISFTLITVFPAGQDQKTWTQFIKWGEIFVDTPLTESIKQLQT
LEGFROILOQSYNTALDDWRKLRLOAPGLPSSALQQAALTAKIRFENVHNDFREIP
GFQLEFYKTLLEPIYACAAANFHLNLOQGAELADEWNADHPQOIEPNAGTSDDYKIL
LKENIPKSYNCANTYREGLNKLNNENRWSIFENDYRYMTIVTLTIAQFSYDIK
RYKDSIRIGGIKTELTREIYTTTEINFDRLTLEIYQNLAIMETNLRSGURLFSFLD
ELIFYKNETYGNRLVGNRSTVATCTEIIYBERTGPTTKTLIPFESYKSVIV
TDQVPTSPFPNIYFTINQIIELYLNNSPSNKLTSAGNLSNDKTKTDFQFPVKDC
KPIINPCLPSSYNSHILSQFSLEYSYKIGLALNLTLYCALCTHSSVNRNNAISD
KITMPEIAKNSLDTNSKVIIEGPHGTGNNLVLSQSGRLEITCTENSTOSYVRLR
YATNAGNTLPNLSLTIPGVIGIPFPQRLNFTSGTNTNNLOYGDFGFPFSTVTLPL
NRNPIPFNRADYSNLSIIIDKIEFIPITSSVRQNRKQKLETTIQTINTPTNHTKN
TLNIEATNYDID"
ORIGIN
Alignment Scores:
Pred. No.: 3,51e-225 Length: 2085
Score: 3104.50 Matches: 598
Percent Similarity: 91.95% Conservative: 30
Best Local Similarity: 87.55% Mismatches: 52
Query Match: 80.24% Indels: 3
DB: 6 Gaps: 2
US-10-783-417-2 (1-735) x CQ868309 (1-2085)
Qy 52 MetCysGlnGlyAsnThrGlnTyrGlyAspAsnPheGluThrPheAlaSerAlaAspThr 71
Db 1 ATGTGTCAGGGAATACACAATATGCTGATATATTTCCAGACATTTGCTAGTCTATACA 60
Qy 72 IleAlaAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAlaGlyIleGlyGly 91
Db 61 ATGTCTCAGTTAGTCAGGCTACTATTGTATCCGCTACTCTGTTAGCCGTATAGGTGG 120
Qy 92 LeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIleSerPheGlyThr 111
Db 121 CTCACCTTCTATATCCGACCGCATAGGAATAATAGTGTCTATAATAATATCTTTTGGTACC 180

Qy 112 LeuileThrValPheTTPProAlaGlyGluGlnAspLysThrValTTPThrGlnPheile 131
Db 181 CTAATCACTGCTTTGGCCCGCGGAGAACAGACAAAACAGATGACAAATTTAT 240
Qy 132 LysMetGlyGluilePheValAspThrProLeuThrGluSerileLysGlnLeuLysLeu 151
Db 241 AAAATGGGAGAAATTTTGTGTGATACACCGTTAAACAGAAACATAAACACAGCTAAAGTTA 300
Qy 152 GlnThrLeuGluGlyPheArgGlnileLeuGlnSerTyrAsnThrAlaLeuAspAspTrp 171
Db 301 CAAACTTTAGAGGATTTAGACAAATATATACAAAGCTATATACAGCATTTAGATGATTGG 360
Qy 172 ArgLysLeuLysArgLeuGlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAla 191
Db 361 AGAAAATTAAGACATACAGCTCCCTGGATTACACCATCATCAGCATTTACACAGCT 420
Qy 192 AlaLeuThrLeuLysileArgPheGluAsnValHisAsnAspPheileArgGluilePro 211
Db 421 GCCTTGACTCTTAAATACGATTTGAGATGTTTCAATGATTTTATTCGAGAAATACCT 480
Qy 212 GlyPheGlnLeuGluThrTyrLysThrLeuLeuLeuProLileTyrAlaGlnAlaAlaAsn 231
Db 481 GGTTCCAACTTGAACTTATAAAGCTATTACTACCTATTATTCGCAAGCTGCTAAT 540
Qy 232 PheHisLeuAsnLeuLeuGlnGlyAlaGluLeuAlaAspGluTTPAsnAlaAspIle 251
Db 541 TTTCAATTAATTTATTACAAACAGGTGCTGAATTTGGCTGATGAATGGAATGCGAGATATA 600
Qy 252 HisProSerGlnileGluProAsnAlaGlyThrSerAspTyrTyrLysLeuLysLeuLys 271
Db 601 CATCTTCACAAATTTGAACCTTATGCTGGAACATCAGATGACTATTATTAACCTTTTAAAA 660
Qy 272 GluAsnileProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsn 291
Db 661 GAAATATACCTTAAATATAGTAATCTTGTGCAAAATACCTATAGAGAAGGACTAAATAA 720
Qy 292 LeuArgAspGluProAsnMetLysTTPSerilePheAsnAspTyrArgTyrMetThr 311
Db 721 CTTGCAAAACCAACCTTATATAGATGAGATATTTAATGATTTATCGAAGATATATGACT 780
Qy 312 IleThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAsp 331
Db 781 ATTACTGTATAGATATATGCTCAATTTCTTTTATGATATTAAGAGATACAAAGAT 840
Qy 332 SerileGlyGlyileGluValLysGlyileLysAsnGluLeuThrArgGluileTyrThr 351
Db 841 TCAATAGSA-----AGNATAGGTGGCTAAAACTGAACCTTACAGAGAAATTTATACA 894
Qy 352 ThrGluileAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMet 371
Db 895 ACTGAAATAAATTTTGACCGCTTACTTACCTTGAAATTTCAACCCAATTCGCTATATATG 954
Qy 372 GluTyrAsnLeuThrArgAlaSerPheLysLysPheSerPheLeuGluGlnPheilePhe 391
Db 955 GAATATATATTAACACGCTTACGGGCTTAGATATTTTCATTTTATAGATGAACCTTATATTT 1014
Qy 392 TyrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyileSerAsnArgAspAlaPro 411
Db 1015 TATACAAAATAGAACTGACGGGAATCGTTAGTTGGATTCGAATTCGTAATAGATCT 1074
Qy 412 ThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySerProThrThr 431
Db 1075 ACTTATGCTACGACAGCACTGAAATATATATATATGAGAGAAACAGGTCCACCCACAACA 1134
Qy 432 LysThrileArgProPheGluSerTyrLysValSerileValThrAspArgGlnSerPro 451
Db 1135 AAAACCTTTAATACCATTTGATCTCTATAAAGCTTTCAATTTGTAATGATAGACAGTAACT 1194
Qy 452 ProValSerProIle---GlnProHisPheileAsnGlnileGluLeuTyrLeuAsn 470
Db 1195 CCTACTTCCCTTTTCCCTAACATATATCTTCAATTAATCAATGAACTTTATTTAAT 1254
Qy 471 GlySerSerAsnAsnThrLeuLysTyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsn 490

Db 1255 AATTCACCTAGTAAATAAATTAACATATTTCAGCTGGGGGAATTTATCTDAATGATAAAAA 1314
Qy 491 ThrThrPhePheGlnPheProArgLysLysAspCysAsnLeuValileAspProGlyCys 510
Db 1315 ACAACTGATTTTCAATTTCTGTAAAAAAGACTGTAAACCAATTTATTAATCCAAATGTT 1374
Qy 511 SerProAsnPheAsnAsnTyrSerHisileLeuSerHisPheSerLeuPheThrTyrSer 530
Db 1375 TTACCAAGCTATAATAGTATAGTCATATTTTATCCAGTTTCTTTATTAATTTATTC 1434
Qy 531 TyrValileGlyLeuGlnLeuGlnileLeuAspThrGlyValileGlyTTPThrHisSer 550
Db 1435 TATAAATTTGATTTAGCGCTAAATATATATATATACAGGTGCATTTAGATGGACACAGT 1494
Qy 551 SerValAspArgTyrAsnAlaileSerAspLysileileThrMetileProAlaileLys 570
Db 1495 AGTGTTAATAGAAATAATGCAATATCAGATAAAATAATTAACAATGATCCAGCAATCAA 1554
Qy 571 GlyAsnAsnLeuAspThrAsnSerLysValileGluGlyProGlyHisThrGlyGlyAsn 590
Db 1555 GGTAAACAGCTCTTGATACAAACTCTAAGGTAATTTGAAGGACCTGGTCATACAGAGGAAAC 1614
Qy 591 LeuValTyrLeuGlnSerGlnGlyArgLeuGluileThrCysGluThrProAsnSerThr 610
Db 1615 TTGGTTTATTTTACAAAGTCAAGGCGTTTATAGAGATTACATGTAGAACTCTTAATTTCTACA 1674
Qy 611 GlnSerTyrPheileArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsn 630
Db 1675 CAATCTTATTACATTTAGACTTCGATACGCTACAAATGGTCTGGAAATATCTTCTCTAAT 1734
Qy 631 IleSerLeuThrIleProGlyValileGlyileProProGlnArgLeuAsnAsnThrPhe 650
Db 1735 ATATCTCTTACAAATACAGGAGTAAATAGGAATACCACCTCAACGACTCAACAAACACTTTT 1794
Qy 651 SerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSer 670
Db 1795 TCTGGTACAAATTAATAATTTTACAAATACGAGATTTTGGGTATTTTCCAAATTTCCAAAT 1854
Qy 671 ThrValThrLeuProLeuAsnArgAsnileProPheilePheAsnArgAlaAspValSer 690
Db 1855 ACAGTAACTTACCTTTTAAATCGAAACATACCAATTTATTTAATCGTGACATGTATCA 1914
Qy 691 AsnSerileLeuileAspLysileGluPheileProLileThrSerSerMethisGln 710
Db 1915 AATTCAAATTTTAAATCAATGATAAAATTTGAAATTTATACCAATTTACTTCTCTGTACGCCAA 1974
Qy 711 AsnArgGluLysGlnLysLeuGluThrileGlnThrLysileAsnThrPhePheThrAsn 730
Db 1975 AATAGAGAAAAACAAAAATTTAGAACTATCCAAACAAAAATAAATACATTTTTTTCACAAAT 2034
Qy 731 HisThrLys 733
Db 2035 CATACAAA 2043
RESULT 4
BTTOXD1
LOCUS 3543 bp DNA linear BCT 18-APR-2005
DEFINITION Bacillus thuringiensis gene for 130 kDa delta-endotoxin.
ACCESSION Y00423
VERSION Y00423.1 GI:40351
KEYWORDS delta-endotoxin; endotoxin.
SOURCE Bacillus thuringiensis
ORGANISM Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE 1 (bases 1 to 3543)
AUTHORS Ward, E.S. and Ellar, D.J.
TITLE Nucleotide sequence of a Bacillus thuringiensis var. israelensis gene encoding a 130 kDa delta-endotoxin
JOURNAL Nucleic Acids Res. 15 (17), 7195 (1987)
PUBMED 2821500
REFERENCE 2 (bases 1 to 3543)

AUTHORS Ward,E.S.
TITLE Direct Submission
JOURNAL Submitted (11-MAY-1987) Ward E.S., University of Cambridge,
Department of Biochemistry, University of Capbridge, Tennis Court
Rd., Cambridge CB2 1OW
COMMENT *strain= var.israelensis;
Data kindly reviewed (12-JAN-1988) by Ward E.S.
FEATURES Location/Qualifiers
source 1. .3543
/organism="Bacillus thuringiensis"
/mol_type="genomic DNA"
/db_xref="taxon:1428"
1. .3543
/notes="unnamed protein product; endotoxin (AA 1-1180)"
CDS
/codon_start=1
/transl_table=11
/protein_id="CAA68485.1"
/db_xref="GI:40352"
/db_xref="GOA:P16480"
/translation="MNPYQNKVEYELNLSAOKKLNISNNYTRYPIENSFKQLQSTNY
KWLNNCOQOQYGGDFETFDISGELSAVTIVVGTVLTGRGETPLGLALIGFTLIP
VLFPQAOQNTWSDFTQTKNIIKKEIASYISNANKILNRSFNVLSTVHNHLKWHN
NFPNQTVQRTQIVLHHPQNVIPFELVNSCPNEDSCDDYINILVLSSTAQAANLHL
TPDSNLGNINWNTYRTKMTTAVLDVVALFPNVDYKYPGVQSELTRIYOVL
NFEESPKYVDFOYQDSLTRRPHLFTWLDLSNLFYEKAOTTPNNFTSHVNMPHYLD
NLSQSSVFGNHNVTDLKSLGLATNIIYFLNVI SLDNKYLNDYNNISKMDFFITNG
TRLLEKELTAGSQIYDVNVNIFGLPILKRNQGNPTLPFYDNYSHLSFKSL
IPATYQVYTAWTHSSDPKNTIYTLTTOIPAVKANSLGTSKVVQPGHGTGDL
IDFKHFKTKQHSNQSYFIRIRVASNSANTRAVINLSIPGVAELGNALNPTFSG
TNYLKYKDFQYLFESNEVFAPNQNISLVFNRSVYNTNITVLIDKIEFLPTIRSR
EDREKQLETVOQIINTFVAPNTKNTLOSELTDYDIOAANLVECTSEELYPKRMLL
LDENVKAKQLSQRNVQLQDPEASLTGWTSDNIIQEDDPPIFKGHYLMHSGARDID
GITFYPIYIQQDESKEKPTPLRVFGSSQVELVSVRGEEDLVSRIEEDLVSRIE
LYPSTDGSCNRCETSPANIGNTSDMLYSCQYDTGKHKVVCQDSHOFSTFDIGA
LDTNENIGVWVFKISPGDQASLDNLEIEEGPIDGEALSRVHKMKWNDQMEAR
SETQOAYDVAKQARLYLDVDEALQFDTLQAOIQAQVYLVQSIIPVYNDWLSVPG
MNYDI VVELDAQARLYLDIRNIIKNGDFTQGVGHVHTGNADVQOQIDGVSLVLS
NWSAGVSNVHLNHHGVYLVIAKKEGPGNGYVTLMDNEENQEKLTFTSCBEGYITK
TVDFPDDTDRVRIEIGETSGSFYIESIELICMNE"
919
old_sequence
/notes="g was a in [1]"
922
old_sequence
/notes="g was a in [1]"
/citation={2}
ORIGIN
Alignment Scores:
Pred. No.: 2,57e-70 Length: 3543
Score: 1061.50 Matches: 271
Percent Similarity: 50.39% Conservative: 118
Best Local Similarity: 35.10% Mismatches: 278
Query Match: 27.44% Indels: 105
DB: 1 Gaps: 24
US-10-783-417-2 (1-735) x BTTOXD1 (1-3543)
QY 1 MetAsnGlnAsnAspAsnGluTyrGluIleAspSerHisThrSerProTyr 20
Db 1 ATGAATCCTTATCAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 60
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 61 -----AATATATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 114
QY 41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
Db 115 TTCAAAAGTACAAATTAATAAAGTGTCTCAATATGTCACAGAAATCAGCAGATGGT 174
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIle 80
Db 1168 ACACCTTGATAATATATATCCCAAAATCTAGTGTGTTTGGAAATACAAATGTAACGTATAA 1227

Db 175 GGAGATTTTGAACCTTTTATTGTAGATG-----GGTGAACCTCAGTGCCTATCTACTATT 225
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
Db 226 GTAGTTGGGACCGTACTGCTGTTTGGGTTTCAACA-----CCCTTAGGA 273
QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120
Db 274 CTT-----GCTTTAATAGGTTTTTGTACATTAATACAGTTCTTTTCCAGCCCA 324
QY 121 GluGlnAspLysThrValThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db 325 GACCAATCTAACACA---TGGAGTGACTTTATACACAACTAAAAATATATTAATAA 381
QY 141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
Db 382 GAAATAGCATCAACATATATATAAGTAAATGCTAATAAATTTTAAACAGGTCGTTTAA 441
QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaPro 180
Db 442 ATCAGCACTTATCATATACCTTAAACATGGAG-----AATAATCCA 486
QY 181 GlyLeuProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGlu 200
Db 487 AACCCCAAAATACTCAGGATGAAGGACACAAATCCAGCTAGTTTACATTTTCAA 546
QY 201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr 218
Db 547 AATGTCATCCAGAGCTTGTAACCTCTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 606
QY 219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGln 238
Db 607 AACATACCTAGTATTATCTAGTTATGCAACAGCAGCAAACTTACATCTGACTGTATAAT 666
QY 239 GlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluPro 258
Db 667 CAAGCCGCTCAAAATTTGAAGCGTATTTAAAAACAATCGACAATTCGATTTATTAGAGCT 726
QY 259 AsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSer 278
Db 727 ---TTGCCAACAGCAATGTAATTTATTCAGTATTGACTAAAGCTATAGAGATATACACT 783
QY 279 AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro----- 296
Db 784 AATTATGTGTACAACCTTATAAAGAGGATTAATTTAATTAACAGCAGCCTGATAGT 843
QY 297 -----AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIle 312
Db 844 AATCTTGATGGAAATATAAATCTGGAACACATACATACGATATCGAACAATAATGACTACT 903
QY 313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer 332
Db 904 GCTGTATTAGATGTTGTGCACTCTTTCCTAATTAATATGATGATAGGTAATAAT- 954
QY 333 IleGlyGlyIleGluValLysGlyIleLysAsnGlnLeuThrArgGluIleTyrThr 352
Db 955 -----CCAATAGTGTCTCAATCTGAACCTTACTCGAGAAATTTATCAGGTA 999
QY 353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
Db 1000 ---CTTAACCTTCAAGAAAGCCCTCTATAAATATTTAT-----GACCTTCAATATCAAGAG 1050
QY 373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyr 392
Db 1051 GATTCACTTACAGTAGACCG---CATTTATTTACTTGGCTTGATCTTTGAAATTTTAT 1107
QY 393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
Db 1108 GAAAAACCGCAAACTACTCTCTAATAATTTTTCACGACCATTAATATGTTTCATTAC 1167
QY 402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
Db 1168 ACACCTTGATAATATATCCCAAAATCTAGTGTGTTTGGAAATACAAATGTAACGTATAA 1227

Qy	421	LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr	440
Db	1228	TTA-----AAATCTCTGGTTGGCAACAATATTTATATT-----	1263
Qy	441	LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe	460
Db	1264	-----	1266
Qy	461	IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr---	476
Db	1267	TTATTAAATGTCATGAAGCTTAGATAATAATAATATCTAAATGATTATAATAATTATTAGTAA	1326
Qy	477	-----LeuLysTyrSerAlaGlyGly	483
Db	1327	ATGGATTTTATTATACTAATGTTACTAGACTTTTGGAGAAGAACTTACAGCAGATCT	1386
Qy	484	SerLeuSerAsnTyrGln---AsnThrThrPheGlnPheProArgLysLysAspCys	502
Db	1387	GGGCAATAAATCTTATGATGATAATAATAATAATTTTCGGGTTACCAATCTTTAAACGAAGA	1446
Qy	503	AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer	522
Db	1447	GAGAAATCAAGGAACCCCTCTTTTCCACATATGATACTATAGTCATATTTATCA	1506
Qy	523	HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuIleLeuAspThr	542
Db	1507	TTTATTAAAGTCTTAGTATCCCTGGCAACATATAAACTCAAGTGTAT-----	1554
Qy	543	GlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle	562
Db	1555	---ACGTTTCTGGACACACTCTAGTGTGATCTTAATAATACATTTATACACATTTA	1611
Qy	563	IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu	582
Db	1612	ACTACCCAAATTCAGCTGTAAAGCGAATTCACCTGGGACTGCTTCTAAGGTTGTTCAA	1671
Qy	583	GlyProGlyHisThrGlyLysAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle	602
Db	1672	GGACCTGGTCATACAGAGGGGATTTAATT-----GATTTCAAAGATCATTTTCAAAAT	1725
Qy	603	ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn	622
Db	1726	ACATGTCACACTCAAAATTTTCAACAATCGTATTTTATAAGAACTCGTTATGCTTCAAT	1785
Qy	623	GlyValaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro	642
Db	1786	GGAAGCGCAATACTCGAGCTGTATATAAATCTTAGTATCCAGGGGTAGCAGAACTG---	1842
Qy	643	ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp	662
Db	1843	GGTATGGCACTCAACCCCACTTTTCTGTGTACAGATTATACGAATTTAAATAATAAAGAT	1902
Qy	663	PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe	682
Db	1903	TTTCAGTACTTGAATTTTCTAAGAGAGGTGAATTTGCTCAAAATCAAAATATCTCTT	1962
Qy	683	IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleAspLysIleGluPhe	701
Db	1963	GTGTTTAATCGTTCGGATGTATATACAAACACACAGTACTTATGATAAATTTGAATTT	2022
Qy	702	IleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGln	721
Db	2023	CTGCCAATTACTCGTTCTATAAGAGAGAGATAGAGAGAAACAAATTTAGAAACAGTACAA	2082
Qy	722	ThrLysIleAsnThrPhePheThrAsnHisThrLys	733
Db	2083	CAAAATAATTAATACATTTTATGCAAAATCCTATAAAA	2118
RESULT 5			
108083			
LOCUS			
DEFINITION Sequence 1 from Patent EP 0296870.			

ACCESSION	I08083		
VERSION	I08083.1	GI:589204	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 4451)		
AUTHORS	Ellar,D.J. and Ward,E.S.		
TITLE	New toxin-encoding DNA fragments from <i>Bacillus thuringiensis</i>		
JOURNAL	subsp. <i>israelensis</i>		
FEATURES	Patent: EP 0296870-A1 1 28-DEC-1988;		
	Location/Qualifiers		
	1..4451		
	source	/organism="unknown"	
		/mol_type="unassigned DNA"	
ORIGIN			
Alignment Scores:		Length: 4451	
Pred. No.:	3,426-70	Matches: 271	
Score:	1061.50	Conservative: 118	
Percent Similarity:	50.39%	Mismatches: 278	
Best Local Similarity:	35.10%	Indels: 105	
Query Match:	27.44%	Gaps: 24	
DB:	6		
US-10-783-417-2 (1-735) x I08083 (1-4451)			
Qy	1	MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr	20
Db	891	ATGAATCCTTCAAAATAAATAAATGAATATGAACATTAATGAATGCTTCACAAAAAATTA	950
Qy	21	PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro	40
Db	951	-----AATATATCTCAATAATATATCAAGATATCAATAGAAAAATAGTCCAAAAACAATTA	1004
Qy	41	LeuGlnAsnThrAsnTyrLysGluThrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly	60
Db	1005	TTACAAAGTACAAATATATAAGATTGGCTCAATATGTGTCAACAGAAATCAGCAGTATGTT	1064
Qy	61	AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle	80
Db	1065	GGAGATTTTGAACATTTTATTGATAGT-----GGTGAACCTCAGTCCCTATATATT	1115
Qy	81	ValSerGlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGly	100
Db	1116	GTAGTTGGGACCGTACTGACTGGTTTCGGGTTCAACA-----CCCTTAGGA	1163
Qy	101	IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly	120
Db	1164	CTT-----GCTTTAATAGTGTTCGTATCAATTAATACCAATTCCTTTTCCAGCCCAA	1214
Qy	121	GluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr	140
Db	1215	GACCAATCAACACA---TGGAGTGACTTTTAAACACAAACTAAATAATATTAAAAAA	1271
Qy	141	ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle	160
Db	1272	GAATAGCATCAACATATATAAGTAATCTAATAAAATTTTAAACAGGTCGTTTAAATGTT	1331
Qy	161	LeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaPro	180
Db	1332	ATCAGCATTATCATCAATCACTTAAACATCGGGAG-----AATAATCCA	1376
Qy	181	GlyLeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGlu	200
Db	1377	AACCCAAATACTCAGGATGTAAAGGACACAAATCCAGTAGTAGTTTACCATTTTCAA	1436
Qy	201	AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr	218
Db	1437	AATGTCATTCCAGAGCTGTAAACTCTGTCTCTCTAATCTAGTAGTATGCGATTACTAT	1496
Qy	219	LysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGln	238

Db 1497 AACATACTAGTATTCTCTAGTTATGACAAAGCAGCAAACTTACATCTGACTGATTAAAT 1556
Qy 239 GlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluPro 258
Db 1557 CAAGCGCTCAAAATTTGAAGCGTATTTAAAAACAATCGCAATTCGATATTATTAGAGCCT 1616
Qy 259 AsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSer 278
Db 1617 ---TTGCCAACAGCAATGATTATTCAGTATTGACTAAAGCTATAGAAGATTACACT 1673
Qy 279 AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro----- 296
Db 1674 AATTATTGTGTACAACACTTATAAAAAAGGATTAATTTAATTAACAGCGCTGATAGT 1733
Qy 297 -----AsnMetLysTspSerIlePheAsnAspTyrArgArgTyrMetThrIle 312
Db 1734 AATCTTGATGGAAATATAACTGGAACACATACATACGTATCGAACAAAAATGACTACT 1793
Qy 313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer 332
Db 1794 GCTGTATTAGATGTTCTGTCACCTCTTCTCTAATTATGATGTAGTAATAT----- 1844
Qy 333 IleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThr 352
Db 1845 -----CCAAATAGGTGTCCTAATCTGAACCTTACTCGAGAAATTTATCAGGTA 1889
Qy 353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
Db 1890 ---CTTAACCTCGAAGAAAGCCCTATAATATAT-----GACTTTCATATCAGAG 1940
Qy 373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyr 392
Db 1941 GATTCACCTACAGTAGACCG---CATTTATTTACTTGGCTTGATCTCTTGAATTTTAT 1997
Qy 393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
Db 1998 GAAAAAGCGCAAACTACTCTCTAATAATTTTTCACCGCCATTATAATATGTTTCATTAC 2057
Qy 402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
Db 2058 ACACCTTGATAATATATCCCAAAATCTAGTGTGTTTGGAAATCACAATGTAACGTGATAA 2117
Qy 421 LeuTyrGlyGluArgThrGlySerProThrLysThrIleArgProPheGluSerTyr 440
Db 2118 TTA-----AAATCTCTGTTGGCAACAATTTATAT----- 2153
Qy 441 LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
Db 2154 ----- 2156
Qy 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
Db 2157 TTATTAATGTCATAGCTTAGATAATAATAATCTAAATGATTATAATATATTAGTAAA 2216
Qy 477 -----LeuLysTyrSerAlaGlyGly 483
Db 2217 ATGGATTTTTTATACTAATGTTACTAGACTTTTGGAGAAAGAACTTACAGCGAGTCT 2276
Qy 484 SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysLysAspCys 502
Db 2277 GGGCAATAACTTATGATGTAATAATAAAATATTTTCGGGTATACCAATCTTAAACGAAGA 2336
Qy 503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
Db 2337 GAGAATCAGGAACCCCTACCCCTTTTCCCAACATATGATACTATAGTATATTTTATCA 2396
Qy 523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnGlnIleLeuAspThr 542
Db 2397 TTTATTAAAAAGTCTTGATATCCCTGCAACATATAAACTCAAGTGAT----- 2444
Qy 543 GlyValLeuGlyTrpThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
Db 2445 ---ACGTTTCTGGACACACTCTAGTGTGATCCTAAAAATAACAATTTATACACATTTA 2501

Qy 563 IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582
Db 2502 ACTACCCAAATTCAGCTGTAAAAGCGAATTCACCTGGGACTGCTTCTTAAGGTTGTTCAA 2561
Qy 583 GlyProGlyHisThrGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle 602
Db 2562 GGAACCTGTGTCATACAGAGGGGATTTAAT-----GATTTCAAGATCATTTCAAAT 2615
Qy 603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
Db 2616 ACATGTCACACTCAAAATTTTCACAATCGTATTTTATAAGAAATCGTTATGCTTCAAT 2675
Qy 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyLeuPro 642
Db 2676 GGAAGCGCAAACTACTCGAGCTGTTTATAAATCTTAGTATCCCGGGTAGCAGAACTG-- 2732
Qy 643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
Db 2733 GGTATGGCACTCAACCCCACTTTTCTGTGTACAGATTATACGAATTTAAAAATATAAAG 2792
Qy 663 PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnAsnArgAsnIleProPhe 682
Db 2793 TTTCACTAGTACATGAAATTTCTAAGAGGTGAAATTTGCTCCAAATCAAAACATATCTCT 2852
Qy 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleIleAspLysIleGluPhe 701
Db 2853 GTGTTTAAATCGTTCGGATGTATATACAAACACACAGTACTTATGATAAAATGAAAT 2912
Qy 702 IleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGln 721
Db 2913 CTGCCAATTTACTCGTTCTATAAGAGGATAGAGAAACAAAAATTAGAAACACGTACAA 2972
Qy 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db 2973 CAAATAATTAATACATTTTATGCAATTCCTATAAAA 3008
RESULT 6
I09103 LOCUS 4451 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 1 from Patent WO 8810305.
ACCESSION I09103
VERSION I09103.1 GI:588188
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4451)
AUTHORS Ellar,D.J. and Ward,E.S.
JOURNAL Patent: WO 8810305-A 1 29-DEC-1988;
FEATURES Location/Qualifiers
source 1..4451
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 3.42e-70 Length: 4451
Score: 1061.50 Matches: 271
Percent Similarity: 50.39% Conservative: 118
Best Local Similarity: 35.10% Mismatches: 278
Query Match: 27.44% Indels: 105
DB: 6 Gaps: 24
US-10-783-417-2 (1-735) x I09103 (1-4451)
Qy 1 MetAsnGlnAsnAsnAspAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 891 ATGAATCCTTATCAAAATAAATAAATGAATATGAACATTAATGCTTCACAAAAAATTA 950
Qy 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 951 -----AATATATCTAATAATATATACAAGATATCCAATAGAAAATAGTCCAAAAACAATTA 1004

Qy 41 LeuGlnAsnThrAsnTyrLysGluThrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
Db 1005 TTAACAAGTACAATTAATAAGATTGGCTCAATATGTCTCAACAGAATCAGCAGTATGGT 1064
Qy 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIle 80
Db 1065 GGAGATTTTGAACCTTTTATGTAGT-----GGTGAACTCAGTGCCTATACTATT 1115
Qy 81 ValSerGlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGly 100
Db 1116 GTAGTTGGGACCGTACTGACTGGTTTCGGGTTCACAACA-----CCCTTAGGA 1163
Qy 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120
Db 1164 CTT-----GCTTTAATAGGTTTGGTACATTAATACAGTCTCTTTTCCAGCCCAA 1214
Qy 121 GluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db 1215 GACCAATCTAACACA---TGGAGTGACTTTTATAACACAACTAAATAATATTATAAAAAA 1271
Qy 141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
Db 1272 GAAATACATCAACATATATAAGTAAGTAACTAAATAAATTTTAAACAGTCTGTTAATGTT 1331
Qy 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaPro 180
Db 1332 ATCAGCACTTATCATATACCTTTAAACATGGGAG-----ATAATCCA 1376
Qy 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGlu 200
Db 1377 AACCCACAAATACTCAGGATGTGAAGACACAAATCCAGCTAGTTCATTACCAATTTCAA 1436
Qy 201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr 218
Db 1437 AATGTCATTCAGAGCTGTGAACCTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1496
Qy 219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGln 238
Db 1497 AACATACTAGTATTATCTAGTTATGTCACAGCAGCAAACTTACATCTGACTGTATTAAAT 1556
Qy 239 GlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluPro 258
Db 1557 CAAGCCGTCAAATTTGAAGCGTATTTAAAAACAACATCGCAATTCGATTTATTAGAGCCT 1616
Qy 259 AsnAlaGlyThrSerAspAspTyrTyrLysLeuLysGluAsnIleProLysTyrSer 278
Db 1617 ---TTGCCAACAGCAATGTATTATCCAGTATTGACTAAAGCTATAGAAATTTACACT 1673
Qy 279 AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro----- 296
Db 1674 AATATTGTGTACAACCTTATAAAAAGGATTAATTTAATTAACAGCCCTGTAGT 1733
Qy 297 -----AsnMetLysTrpSerIlePheAsnAspTyrArgA-gTyrMetThrIle 312
Db 1734 AATCTTGATGGAATATAAACTGGAACACATACATACGTATCGAACAATAATGACTACT 1793
Qy 313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer 332
Db 1794 GCTGTATTAGATGTTGTGCACTCTTTCTTAATATTATGATGTAGTAATAT----- 1844
Qy 333 IleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThr 352
Db 1845 -----CCAATAGGTGTCGAATCTGAACCTTACTCGAGAAATTTATCAGGTA 1889
Qy 353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
Db 1890 ---CTTAACCTTCGAAGAAGCCCTATATAATTTAT-----GACTTTCAATATCAAGAG 1940
Qy 373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyr 392
Db 1941 GATTCACCTACACGTAGACCG---CATTTATTACTTGGCTTGATCTCTTGAATTTTAT 1997

Qy 393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
Db 1998 GAAAAAGCGCAAACTACTCTTAATAATTTTTTCCACGCAATATAATATATGTTTCATTAC 2057
Qy 402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
Db 2058 ACATTGATATATATCCAAAAATCTAGTGTGTTTGGAAATCACAATGTAATCTGATAAA 2117
Qy 421 LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440
Db 2118 TTA-----AAATCTCTGGTTTGGCAACAATAATTTATATT----- 2153
Qy 441 LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
Db 2154 ----- 2156
Qy 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
Db 2157 TTATTAAATGTCAAGCTTAGATAATAAATATCTAAATGATTATAATAATATTAGTAAA 2216
Qy 477 -----LeuLysTyrSerAlaGlyGly 483
Db 2217 ATGGATTTTTTATAACTAATGGTACTAGACTTTTGGAGAGAACTTACACAGGATCT 2276
Qy 484 SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysLysAspCys 502
Db 2277 GGGCAAAATAACTATGATGATAAATAAATAATTTTCGGGTTACCAATCTTTAAACGAAGA 2336
Qy 503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
Db 2337 GAGAAATCAAGAAACCCCTACCCCTTTTCCACATATGATAACTATAGTCATATTTATCA 2396
Qy 523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThr 542
Db 2397 TTTATTTAAAGTCTTAGTATCCCTCGACACATATAAACTCAAGTCTAT----- 2444
Qy 543 GlyValLeuGlyTrpThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
Db 2445 ---ACGTTTGTGGACACACTCTAGTGTGATCTTAAATAATACAAATTTATACACATTTA 2501
Qy 563 IleThrMetIleProAlaIleLysGlyAsnLeuAspThrAsnSerLysValIleGlu 582
Db 2502 ACTACCCAAATTTCCAGCTGTAAAGCGAAATTCACCTGGGACTCTTCAAGGTTGTTCAA 2561
Qy 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle 602
Db 2562 GGACTGTGTACATACAGGCGGATTAAT-----GATTTCAAGATCATTTCAAAAT 2615
Qy 603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
Db 2616 ACATGTCAACACTCAAATTTTCAACAATCGTATTTTATAAGAAATTCGTATGCTTCAAAAT 2675
Qy 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db 2676 GGAAGCGCAATATCTCGAGCTGTGTATAAATCTTAGTATCCCAAGGGGTAGCAAACTG--- 2732
Qy 643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
Db 2733 GGTATGGCACTCAACCCACCTTTCTGTCAGATTAACGANTTAAATATAAAGAT 2792
Qy 663 PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
Db 2793 TTTTCAGTACTTGAATTTTCTAACGAGGTGAAATTTGCTCCAAATCAAAACATATCTCTT 2852
Qy 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleIleAspLysIleGluPhe 701
Db 2853 GTGTTTAACTCGTTCGGATGTATACAAACACACACAGTACTTATTGATAAAATTTGAATTT 2912
Qy 702 IleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGln 721
Db 2913 CTGCCAATTTACTCGTTCATTAAGAGGATGAGAGAAACAAATAATTAGAACAGTACAA 2972
Qy 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733


```

Db          2973 CAAATAAATAATACATTTTATGCAAAATCCTATAAA 3008
|||||
RESULT 7
BACISR44
LOCUS      BACISR44             4253 bp      DNA          linear      BCT 17-FEB-1998
DEFINITION Bacillus thuringiensis israelensis plasmid gene for 130 kDa
            insecticidal protein (ISR4), complete cds.
ACCESSION D00248
VERSION    D00248.1  GI:216289
KEYWORDS   130 kDa insecticidal protein (ISR4).
SOURCE     Bacillus thuringiensis serovar israelensis
            Bacillus thuringiensis serovar israelensis
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
            cereus group.
REFERENCE  1 (bases 1 to 4253)
AUTHORS    Sen, K., Honda, G., Koyama, N., Nishida, M., Neki, A., Sakai, H.,
            Himeno, M. and Komano, T.
TITLE      Cloning and nucleotide sequences of the two 130 kDa insecticidal
            protein genes of Bacillus thuringiensis var. israelensis
JOURNAL    Agric. Biol. Chem. 52, 873-878 (1988)
COMMENT    The genes of ISR3 and ISR4 that were 130 kDa insecticidal
            proteins of BTI were sequenced and compared with other insecticidal
            protein genes. ISR3 and ISR4 were identical in a region of the
            C-terminal 467 amino acids.
FEATURES   Location/Qualifiers
            1..4253
                /organism="Bacillus thuringiensis serovar israelensis"
                /mol_type="genomic DNA"
                /strain="HD522"
                /sub_species="israelensis"
                /db_xref="taxon:1430"
                /clone="pBGH4"
                /plasmid="pBTI-5 and pBTI-6"
                /notes="976 bp downstream of HindIII site"
            381..385
                /notes="Shine-Dalgarno sequence"
            393..3935
                /codon_start=1
                /transl_table=11
                /product="130 kDa insecticidal protein (ISR4)"
                /protein_id="BAA00179.1"
                /db_xref="GI:216290"
                /translations="MNPYQKNBYETLNASOKKLINSNTRYPIENSPKOLLQSTNY
                KWLNMCCNQOQGGDPETFDIGSELASATVVGTVLTCGFPTPIGLALIGGTLIP
                LFLPQAQSNQSDFTITQKNIIKKEIASYISNANKILNRSFNVLSTVHNHLKWTEN
                NPNPQTDVQTQIQVHFNQYVPELVNSPCNPFSDCCDYINLVLSVTAQAQNLHJ
                TVLNQAKFAYELKNNRQFDYELPLETADIVLPVTKALIEDYTNVCTVYKKGILIK
                TTVNLDNGINNNYNTYRTKMTATVLDIVLFPNYDVKGVPYQVSELTRIEYOVL
                NPESPVKYVFOYQESLSTRPLHFTWLDLSNLFYKCAOTTNNFTSHVNMPHYTLD
                NISOKSSVFGNHNVTDKLSLGTATNIYIPLNVI SLDNKYLNDYNI SKMDPFITNG
                TRLLEKELTAGSQIYIDVNVKNIFGLPKRRENQGNPILFPYDNYSHLISFKLSLS
                IPATYQIVTTFATWSVDPRKNTIYTHLTQIPAVKANSLGTASKVQVQPGHTGGDL
                IDPKDFHKITQHSNQQSYFTRIRVASGNSANTRAVINLSIPGVAELGALNLPFTSG
                TDYNLKXFDQYLFESNEVFAPNQNIQSLVFNRSVYNTVTLIDIKLIFLPTIRIR
                EDRKQKLETVQOIINTFVAPNTKNTQISLTDYIDQAAVLVECI SEELYPKEMLL
                LDENVKAKOLSQSRNLVQNGDFESATLGTWTSNTITQEDDPIFKGHYLMHMGADID
                GTIPFYIYIQKIDESKLKYPHTRVGVGSSKDKVLSVSVRSGEEDALMNVPADLNY
                LYPTFDTCESNGCETSAVPANIGTNSDMLSYCYDQTKKKHVVCQDSHFQSFITD7GA
                LDTNENIGVWMPKISGPDGYASLDNLEVEEGIDGELASLRVKHMEKKWNDQMEAR
                SETQNGDYAKAKAIDALFYDQVDEALQFDTTLAIQIYA EYLVSQIPYVNDMLSDVPG
                MNYDIYVELDARVAQARLYVDTRNIYKNGDFTQGVMGHVTGNADYVQDIDGVSVLVS
                LKSNAGSVYVHLOHNGHYLVRTAKKEGPGNGYVTLMDCEENQEKLTFTTSCBEGYITK
                TVDVFDDTRVRIEIGSTEGSFIESIELICMNE"
ORIGIN
Alignment Scores:
Pred. No.:      3,85e-70      Length:      4253
Score:          1060.50      Matches:      271
Percent Similarity: 50.39%      Conservative: 118
Best Local Similarity: 35.10%      Mismatches:  278
Query Match:    27.41%      Indels:      105
DB:             1           Gaps:        24

```


1347 Db -----CCAATAGGTGTCCAATCTGAACCTTACTCGAGAAATTTATCAGGTA 1391
353 Qy GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
1392 Db ---CTTAACCTCGAAGAACCCCTATAATATTAT-----GACTTTCAATATCAAGAG 1442
373 Qy TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyr 392
1443 Db GATTCACTTACAGTAGACCG---CATTATTTACTTGGCTTGATCTTTGAAATTTTAT 1499
393 Qy -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
1500 Db GAAAAAGCGCAACTACTCTCTAATAATTTTTCACCGCCATTATAATATGTTTCATTAC 1559
402 Qy ---LeuValGlyIleSerAsnArgAlaProThrTyrSerAsnThrIleThrGluThr 420
1560 Db ACATCTGATATAATATATCCAAAAATCTAGTGTTTTGGAAATCACAATGTAACTGATAAA 1619
421 Qy LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440
1620 Db TTA-----AAATCTCTGTTTGGCAACAATATTATAT----- 1655
441 Qy LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
1656 Db ----- 1658
461 Qy IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnThr--- 476
1659 Db TTATTAATCTCATAGCTTAGATAATAATAATATCTAAATGATTATAATAATATTAGTAAA 1718
477 Qy -----LeuLysTyrSerAlaGlyGly 483
1719 Db ATGGATTTTTTATACTAATAGTACTAGACTTTTGGAGAAAGAACTTACAGCAGGACTCT 1778
484 Qy SerLeuSerAsnTyrGln---AsnThrThrPheGlnPheProArgLysLysAspCys 502
1779 Db GGGCAATAACTTATGATGATAATAATAATAATATTTTCGGTACCATACTTTAAACGAGA 1838
503 Qy AsnLeuValIleAspProGlyCysSerProAsnPheAsnThrSerHisIleLeuSer 522
1839 Db GAGAAATCAAGGAACCCCTACCCCTTTTCCAAATATGATACTATGATCATATTTTATCA 1898
523 Qy HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThr 542
1899 Db TTTATTAAAAAGTCTTAGTATCCCTCGCAACATATATAAACTCAAGTGAT----- 1946
543 Qy GlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
1947 Db ---ACGTTTCTGGACACACTCTAGTGTGTGATCTTAAATAACAATTTATACACATTTA 2003
563 Qy IleThrMetIleProAlaIleLysGlyAsnLeuAspThrAsnSerLysValIleGlu 582
2004 Db ACTACCAATTTCCAGCTGTAAAGCGAATTCACCTGGGACTGCTTCTAAGTTGTCAA 2063
583 Qy GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGlu 602
2064 Db GGACCTGGTCATACAGGAGGGATTTAATT-----GATTCAAGATCATTTCAAAATT 2117
603 Qy ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
2118 Db ACATGTCAACACTCAAAATTTTCAACATCGTATTTTATAAGAAATCGTGTGCTTCAAT 2177
623 Qy GlyValGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyPro 642
2178 Db GGAAGCGCAATACTCGAGCTGTTATTAATCTTAGTATCCAGGGGTAGCAGAACTG--- 2234
643 Qy ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
2235 Db GGTATGSCACTCAACCCCACTTTTCTGTTACAGATTATACGATTTAAATAATAAGAT 2294
663 Qy PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
2295 Db TTTTCAGTACTTAGAATTTTCTAACGAGGTGAAATTTGCTCCAAATCAAAACATATCTCTT 2354

683 Qy lIlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleAspLysIleGluPhe 701
2355 Db GTGTTTAATCGTTCGGATGTATATACAAACACACAGTACTTATTGATAAAATTTGAATTT 2414
702 Qy lIleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGln 721
2415 Db CTGCCAATTTACTCGTTCATTAGAGAGGATAGAGAGAAACAAATTTAGAAACAGTACAA 2474
722 Qy ThrLysIleAsnThrPhePheThrAsnHisThrLys 733
2475 Db CAATAAATTAATACATTTTATGCAAAATCCTATAAAA 2510

RESULT 8
E01676
LOCUS E01676
DEFINITION DNA sequence encoding Bacillus thuringiensis insectivorous protein.
ACCESSION E01676
VERSION E01676.1 GI-2169929
KEYWORDS JP 1988230090-A/1.
SOURCE Bacillus thuringiensis
ORGANISM Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE 1 (bases 1 to 4934)
AUTHORS Komano, T. and Himeno, M.
INSECTICIDAL PROTEIN OF STRAIN BACILLUS THURINGIENSIS VAR
TITLE Patent: JP 1988230090-A 1 26-SEP-1988;
JOURNAL SUMITOMO CHEM CO LTD
COMMENT OS Bacillus thuringiensis
PN JP 1988230090-A/1
PD 26-SEP-1988
PF 19-MAR-1987 JP 1987066844
PI KOMANO TORU, HIMENO MICHIO
PC C12N15/00, A01N63/00, C12N1/20, C12P21/02, C12N1/20, (C12N1/20, C12R1:13);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=HD522;
FH Key Location/Qualifiers
FH 5'UTR 621..1368
FT CDS 1369..4911
FT /product='Bacillus thuringiensis insectivorous protein'.
FT Location/Qualifiers
FT 1..4934
ORIGIN
Alignment Scores:
Pred. No.: 4.63e-70 Length: 4934
Score: 1060.50 Matches: 271
Percent Similarity: 50.39% Conservative: 118
Best Local Similarity: 35.10% Mismatches: 278
Query Match: 27.41% Indels: 105
DB: 6 Gaps: 24
US-10-783-417-2 (1-735) x E01676 (1-4934)
Qy 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleAspSerHisThrSerProTyr 20
Db 1369 ATGAATCTTATCAAAATAAAATGAATATGAATCAATTAATGAATGCTTCACAAAAAAATTA 1428
Qy 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 1429 -----AATATATCTTAATAATTTATACAAAGATATCAATAGAAAATAGTCCAAACCAATTA 1482
Qy 41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60

Db	1483	TTCAAAAGTCAAAATTAAGATTGGCTCAATATGTGTCAACAGAATCAGCAGTATGGT	1541
Qy	61	AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIle	80
Db	1543	GGAGATTTCGAACTTTTATTGTAGT-----GGTGAACTCAGTGCCTATACTATT	1593
Qy	81	ValSerGlyThrLeuLeuAlaGlyIleGlyLeuThrSerGlyProIleGly	100
Db	1594	GTAGTTGGGACCGTACTGACTGGTTTCGGTTCACAACA-----CCCTTAGGA	1641
Qy	101	IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly	120
Db	1642	CTT-----GCTTTAATAGGTTTGGTACATTAATACAGTCTCTTTTCCAGCCCAA	1692
Qy	121	GluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr	140
Db	1693	GACCAATCTAACACA---TGAGTGACTTTATAACACAAACTAATAATATTATAAAAAA	1749
Qy	141	ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle	160
Db	1750	GAATATGATCAACATATATAAGTAAGTAAATAAATTTTAAACAGTCGTTTAAGTT	1809
Qy	161	LeuGlnSerTyrAsnThrAlaLeuAspAspTyrPargLysLeuLysArgLeuGlnAlaPro	180
Db	1810	ATCAGCACTTATCATATACACTTAAACATGGAG-----AATAATCCA	1854
Qy	181	GlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGlu	200
Db	1855	AACCCAAATAACTTCAGGATGTAAGACACAATCCAGCTAGTTCATTACCAATTTCAA	1914
Qy	201	AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr	218
Db	1915	AATGTCATCCAGAGCTGTGAACCTCTGTCTCTTAATCCTAGTGATTGCGATTACTAT	1974
Qy	219	LysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGln	238
Db	1975	AACATACGTAGTATTATCTAGTTATGCAACAGCAGCAAACTTACATCTGACTGATTAAAT	2034
Qy	239	GlnGlyAlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluPro	258
Db	2035	CAGCCGTCAAAATTTGAAGCGTATTTAAAAACAATCGCAATTCGATTTATTAGACCT	2094
Qy	259	AsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSer	278
Db	2095	---TTGCCAACGCAATTTGATTATTCAGTATTGACTAAAGCTATAGAAGATTACACT	2151
Qy	279	AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro-----	296
Db	2152	AATATTGTGTAACTTATAAAAAAGGATTAAATTTAATTAATAACGCGCTGATAGT	2211
Qy	297	-----AsnMetLysTyrSerIlePheAsnAspTyrArgArgTyrMetThrIle	312
Db	2212	AATCTTGATGGAATATAAATCGGAACACATACATACGATACGATCGAACAATAAGACTACT	2271
Qy	313	ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer	332
Db	2272	CTGTATTAGATCTTGTGTGACTCTTCTTAATATGATGATAGTAAATAT-----	2322
Qy	333	IleGlyIleGluValLysIleLysAsnGluLeuThrArgGluIleTyrThrThr	352
Db	2323	-----CCAATAGGTGTCCAAATCTGAACCTTACTCGAGAAATTTATCAGGTA	2367
Qy	353	GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu	372
Db	2368	---CTTAACCTCGAAGAACCCCTATAAATATTAT-----GACTTTCATATCAAGAG	2418
Qy	373	TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnPheIlePheTyr	392
Db	2419	GATTCATCTACCGTAGACCG---CATTTATTACTTGGCTTGATCTTTGAATTTTTAT	2475
Qy	393	-----ThrGluAsnThrAsnPheGlyAsnArg-----	401


```
Db 95478 GATTCACTACAGTAGACCG---CATTATTATTACTTGGCTTGATCTCTTTGAAATTTTAT 95422
Qy 393 -----ThrGluAsnThrAsnGlyAsnArg----- 401
Db 95421 GAAAAAGCGCAAACTACTCTCTAAATATTTTTCACCGCAATTAATAATATGTTTCATTAC 95362
Qy 402 ---LeuValGlyIleSerAsnArgAlaProThrTyrSerAsnThrIleThrGluThr 420
Db 95361 ACATTGATATATATCCCAAAATCTAGTGTGTTTGGAAATCACATGTAACHTGATAA 95302
Qy 421 LeuTyrGlyGluArgThrGlySerProThrThrIleThrIleArgProPheGluSerTyr 440
Db 95301 TTA-----AAATCTCTGTGTTGGCAACAATAATTTATATT----- 95266
Qy 441 LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
Db 95265 ----- 95263
Qy 461 IleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
Db 95262 TTATTAAATGCTAGCTTAGATAATAAATATCTAAATGATTAATAATATTAGTATAA 95203
Qy 477 -----LeuLysTyrSerAlaGlyGly 483
Db 95202 ATGGATTTTTTATACTAATAGTACTAGACTTTTGGAGAAAGAACTACAGCAGATCT 95143
Qy 484 SerLeuSerAsnTyrGln---AsnThrThrPheGlnPheProArgLysLysAspCys 502
Db 95142 GGGCAAAATACTTATGATGTAATAAATAATTTTCGGGTACCAATCTTAAACGAAGA 95083
Qy 503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
Db 95082 GAGAATCAAGGAACCCCTACCTTTTCCAAATATGATACTATAGTACTATATTATCA 95023
Qy 523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnIleLeuAspThr 542
Db 95022 TTTATTAAAGTCTTAGTATCCCTGCAACATATAAACTCAAGTGAT----- 94975
Qy 543 GlyValLeuGlyTrpThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
Db 94974 ---ACGTTTGCTTGGACACACTCTAGTGTGATCCTTAAATAACAATTTATACACATT 94918
Qy 563 IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582
Db 94917 ACTACCAAAATCCAGCTGTAAGAGCAATTCACCTGGGACTGCTTCAAGGTGTTCAA 94858
Qy 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle 602
Db 94857 GGACCTGGTCTATACAGGAGGGGATTTAAAT-----GATTCAAGAGATCATTTCAAAAT 94804
Qy 603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
Db 94803 ACATGTCAACACTCAAAATTTTCAACATCGTATTTATAGAAATCGTTATGCTTCAAT 94744
Qy 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db 94743 GGAAGCGCAAACTACGAGCTGTATATAAATCTTAGTATCCAGGGGTAGCAGAACTG--- 94687
Qy 643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
Db 94686 GGTATGGCACTCAACCCCACTTTTCTCGTACAGATTATACGAATTTAAATAATAAAGAT 94627
Qy 663 PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
Db 94626 TTTCAGTACTTGAATTTCTTAACGAGGTGAATTTTCTGCTCAAAATCAAAACATATCTCTT 94567
Qy 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleAspLysIleGluPhe 701
Db 94566 GTGTTTAATCGTTCGGATGTATATACAAACACACAGACTACTTATTGATAAAATGAATTT 94507
Qy 702 IleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGln 721
Db 94506 CTGCCAATTACTCGTTCATAGAAGGATAGAGAGAAACAAAATTTAGAAACAGTACAA 94447
```

```
Qy 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db 94446 CAAATAATTAATACATTTTATGCAAAATCCTATAAA 94411

RESULT 10
BACMSQB
LOCUS 3753 bp DNA linear BCT 26-APR-1993
DEFINITION B.Churingiensis insecticidal endotoxin gene, complete cd.
ACCESSION M12662
VERSION M12662.1 GI:143228
KEYWORDS
SOURCE
ORGANISM
Bacillus thuringiensis
Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
cereus group.
REFERENCE
1 (bases 1 to 3753)
AUTHORS Thorne, L., Garduno, F., Thompson, T., Decker, D., Zounes, M., Wild, M.,
Walfield, A. M. and Pollock, T. J.
TITLE Structural similarity between the lepidoptera- and diptera-specific
insecticidal endotoxin genes of Bacillus thuringiensis subsp.
'kurstaki' and 'israelensis'
J. Bacteriol. 166 (3), 801-811 (1986)
3011746
JOURNAL
PUBMED
COMMENT Original source text: B.thuringiensis subspecies israelensis DNA.
FEATURES
source
Location/Qualifiers
1..3753
/organism="Bacillus thuringiensis"
/mol_type="genomic DNA"
/db_xref="taxon:1428"
941..2968
/note="insecticidal endotoxin (put.); putative"
/codon_start=1
/transl_table=11
/protein_id="AAA22614.1"
/db_xref="GI:143229"
/translation="MNPQNKNEYEIFNAPNGFSKNYSRYPLANKPNQPLKNTNY
KDLNVQDQNNQYGNAGNFASSETIVGVSGAIIIVGTMGFAFAAFLAAGIISFOTL
LPIFWQSDPANVWQDLNIGRPIQIEIDKNIINVLSIVTPIKQLDKYQBFKWE
PARTHANAKAVHDLFTLEPIIDKDLMLKNASVRIPTLPAYAQIATWHLNLKHA
TYNNIWLQNGINPSTFNSSNYGYGLARKIQEYTDYCIOTYAGLTMLRTNTATWN
MYNTRYLEMTLVLDLAIAPNYDPEKYPGVKSELIREVTNVNSDTPRTITELNG
LSRDTPLTFINQGRFYTRNSRDILDPYDIFSTGNQMAFTHNDNRNLIWGAVHNI
ITNPTQVFPFPPKRPDKVEIRVREYSDIIYEMIFPSNSEVSEVRSNYSNIENYK
RITSYMLPKOTWKNEEYCHTLYKTDNYIFSVRRERVRVAFSWHTSVDFTDILD
NTOIHALKALKVSSDSKI VKGPGHTGGDLVLKDSMDRVRFLKNVSRQYQVRIYA
TNAPKTVFLTGIDTISVELSPSTSRQNPNDLTLYADFGYVTFRTVPNKTFEGBDT
LMLTLTGPNHSNIYIDKIEFIPITQSVLDYTEKQIEKTKIVNDLPVN"
ORIGIN

Alignment Scores:
Pred. No.: 7,08e-58 Length: 3753
Score: 897.50 Matches: 247
Percent Similarity: 48.32% Conservative: 126
Best Local Similarity: 31.99% Mismatches: 270
Query Match: 23.20% Indels: 129
DB: 1 Gaps: 30

US-10-783-417-2 (1-735) x BACMSQB (1-3753)
Qy 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 941 ATGAATCATATCAAAATTAAGAATGAATATGAAATATTCAATGCTCCATCGATGTTT 1000
Qy 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 1001 -----AGCAAGTCTAATAACTATTCTAGATATCCATTAGCAAAATAAGCAATCAACCA 1054
Qy 41 LeuGlnAsnThrAsnTyrLysGluTrpLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
Db 1055 CTGAAAACACGAATTACAAAGATTGGCTCAATGTGTGTCAGAGATAATCAACAATATGC 1114
Qy 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80
```

Db 1115 AATAATCGGGGAATTTTGGTAGTCTCGAAACTATTGTTGAGTTAGTCAGAGTATTATT 1174
 Qy ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
 Db 1175 GTAGTAGAACTATGTAGGA-----GCTTTGCTGCCCT----- 1210
 Qy IleIleGlyAlaIleIleSerPheGlyThrIleuIleThrValPheTyrProAlaGly 120
 Db 1211 GTCTTAGTCGAGGTATAATATCTTTTGGGACTTTGTTGGCATCTCTTTGG---CAAGGA 1267
 Qy GluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
 Db 1268 TCTGACCCCTGCAAAATGTTTGGCAGGATTGTTAAACATCGGA-----GGAAGG 1315
 Qy ProLeuThrGlu-----SerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPhe 157
 Db 1316 CCTATACAGAAATAGATAAACAATTAATTAATGACTAACTTCTATCGTAACACCTATA 1375
 Qy ArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrPargLysLeuLysArgLeu 177
 Db 1376 AAAAAATCAACTTGATAAATATCAAGAATTTTTTCGATAAATGGGAGCCAGCACGT--- 1429
 Qy GlnAlaProGlyLeuProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIle 197
 Db 1430 -----ACACACGCTAAATGCTAAAGCAGTACATGATCTCTTTACT 1468
 Qy ArgPheGluAsnValHisAsn-----AspPheIleArgGluIleProGlyPheGln 214
 Db 1469 ACCTTAGACCTATATAGATAAAGATTAGATATGTTAAAAATAATATGCTAGTATCGA 1528
 Qy LeuGluThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAsnPheHisLeu 234
 Db 1529 ATACCAACA-----CTCCCTGCATATGCACAAATAGTACTTGGCACTTG 1573
 Qy AsnLeuLeuGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSer 254
 Db 1574 AATTATTAAAAATGCTGCTACCTATTACAATATATATGG-----CTGCAAAATCAA 1624
 Qy GlnIleGluProAsnAlaGlyThrSerAspTyrTyrLys---LeuLeuLysGluAsn 273
 Db 1625 GGTATAAATCCAAGTACTTTCAATTATCAATTAATCTATCAGGGCTATTAAACGTAAA 1684
 Qy IleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArg 293
 Db 1685 ATACAAGATATAGTACTATTTGATACAAACGTACAAATCAGGACTAACTATGATTAGA 1744
 Qy AspGluProAsnMetLysTyrSerIlePheAsnAspTyrArgArgTyrMetThrIleThr 313
 Db 1745 ACTAATACTAACGCAACATGGAATATGATAAATCTTACCGTTTGGAAATGACTTAACT 1804
 Qy ValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIle 333
 Db 1805 GTGTTAGATCTTATGCTATTTTCCAAATATGACCCAGAAAATAT----- 1852
 Qy GlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThrGlu 353
 Db 1853 -----CCAAATAGGAGTTAATCTGAACTTATCAGAGAGTTTAT---ACGAAT 1897
 Qy IleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyr 373
 Db 1898 GTTAATAATCAGATACATTT-----AGAACCAATAACAGAACTAGAAAAT 1939
 Qy AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnPheIlePheTyrThr 393
 Db 1940 GGATTAATACTAGA---AATCTCATATTATTCTTGGATAAACCAACGAGCGCTTTTACACA 1996
 Qy GluAsnThrAsn-----PheGlyAsnArgLeu 402
 Db 1997 AGAAATCTCGAGACATCTTGATCCTTATGATATTTTCTTTTACAGGTACACAGATG 2056
 Qy ValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyr 422

2057	Db	-----GCCTTTACACATACTAATGATGATGCCAACATAAATCTCGG	2095
423	QY	GlyGluArgThrGly-----SerProThrThrLysThrIleArgProPheGluSer	439
2096	Db	GGAGCGGTTTCATGGAATATATTTCTCAAGACACATCCAAAGTATTTCTTTTATAGA	2155
440	QY	Tyr-----LysValSerIleValThrAspArgGlnSerProProValSerPro	455
2156	Db	AACAAACCTATTGATAAGGTCGAATTTGTTCAGACATAGAGAGTACTCAGATATAATA	2212
456	QY	IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn	475
2213	Db	-----TATGAAATGATATATTTTTCGAATAGACGAGTAA	2245
476	QY	ThrLeuLysTyrSerAlaGlyGlySerLeu---SerAsnTyrGlnAsnThrThrPhePhe	494
2246	Db	GTATTTTCGATATTCATCCATTCACAAATAGAAAATATATTATAAAGAACTGATCTTAT	2305
495	QY	GlnPheProArgLysLysAspCysAsnAsnLeuValIleAspProGlyCysSerProAsnPhe	514
2306	Db	ATGATTCCAAAACAAACATCGAAAAAT-----	2332
515	QY	AsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGly	534
2333	Db	GAAGAAATATGGTCATCATCTCTATCGTATATAAAACGTAGATAATTATATATTTT	2392
535	QY	LeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArg	554
2393	Db	AGAGAAAGAGACAGATT-----GCATTAGTTGGACACATACTAGTGTGTTGATTC	2443
555	QY	TyrAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeu	574
2444	Db	CAAAATACAAATAGATTTAGATAACATCACCCAAATCCACGCTCTAAAGCTTTGAAGGTA	2503
575	QY	AspThrAsnSerLysValIleGluGlyProGlyHisThrGlyAsnLeuValTyrLeu	594
2504	Db	AGTTCTCATTCGAAAATTTGCAAGGTCCTCGGTCCACAGGTGGAGATTTGGTAATCTTT	2563
595	QY	GlnSerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPhe	614
2564	Db	AAAGATAGTAGTGGATTTTAGAGTTAGATT---TTAAAAAATGTTTCTGCACAAATCA	2620
615	QY	IleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThr	634
2621	Db	GTACGTATTTCGTTATGCTACTAATGCT-----CCAAAGACACACAGATATTC	2665
635	QY	IleProGlyValIleGlyIleProGlnArgLeuAsnAsnThrPheSerGlyThrAsn	654
2666	Db	TTAACCCGGAATAGATACTATAAGTGTGGAG---CTCCCTAGTACCCTCCCGCCAAAC	2722
655	QY	TyrAsn-----AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThrVal	672
2723	Db	CCAATGCTACAGATTTTAAATATGACAGATTTTGGATATGTAACATTTCCAGAACAGTT	2782
673	QY	-----ThrLeuProLeuAsnArgAsnIleProPheIle	683
2783	Db	CCAAATAAAACATTTGAAGGAGAGACACTTATTATATGACC-----TTA	2827
684	QY	PheAsnArgAlaaspValSerAsnSerIleIleAspLysIleGluPheIlePro	703
2828	Db	TATGGTACACCAATCATTCATATAATATA---TATATTGACAAAATCGAAATTTATCCA	2884
704	QY	IleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLys	723
2885	Db	ATCACTCAATCTGTATTAGATTATACAGAGAGCAGCAAAATATAGAAAAAACACAGAAAATA	2944
724	QY	IleAsnThrPhePheThrAsnHisThrLysThrLeu	735
2945	Db	GTGAATGATTTATTTGTTTAATTAATAACAAAGTTCTT	2980
RESULT 11			
BTPTOXIS			
LOCUS			
	BTPTOXIS	127923 bp	DNA linear BCT 16-APR-2005

RESULT 11
BTPBTOXIS
LOCUS

```

DEFINITION Bacillus thuringiensis subsp. israelensis plasmid pBtoxis.
ACCESSION AL731825
VERSION AL731825.1 GI:21685410
KEYWORDS .
SOURCE
ORGANISM
Bacillus thuringiensis serovar israelensis
Bacillus thuringiensis serovar israelensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
cereus group.
REFERENCE
1 (bases 1 to 127923)
AUTHORS Berry,C., O'Neil,S., Ben-Dov,E., Jones,A.F., Murphy,L., Quail,M.A.,
Holden,M.T., Harris,D., Zaritsky,A. and Parkhill,J.
TITLE Complete sequence and organization of pBtoxis, the toxin-coding
plasmid of Bacillus thuringiensis subsp. israelensis
JOURNAL Appl. Environ. Microbiol. 68 (10), 5082-5095 (2002)
PUBMED 12324359
REFERENCE
2 (bases 1 to 127923)
AUTHORS Parkhill,J.
TITLE Direct Submission
JOURNAL Submitted (19-APR-2002) Submitted on behalf of the pBtoxis
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
NOTES:
Details of pBtoxis sequencing at the Sanger Centre are available on
the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/B_thuringiensis/).
FEATURES
source
1..127923
Location/Qualifiers
/organism="Bacillus thuringiensis serovar israelensis"
/mol_type="genomic DNA"
/db_xref="taxon:1430"
/notes="plasmid pBtoxis"
1145..1149
1163..2089
/genes="pBt001"
1163..2089
/genes="pBt001"
/notes="Similar in part to Bacillus anthracis pxol-49
TR:Q9X319 (EMBL:AF065404) (227 aa) fasta scores: E():
8.9e-44, 78.48% id in 158 aa"
/codon_start=1
/transl_table=11
/protein_id="CAD330064.1"
/db_xref="GI:21685411"
/db_xref="UniProt/TREMBL:Q8KNX5"
/translations="MSTHLNFSQNFVKVGVKFDGLFNIILKENERKELPSSALATY
IMLHFCNDIGMLPRFQIMDLAKSGIPVTYITGTCVCLERKLVREIPVGNRTYVE
IVDYALNHTAETMTMINISLSYRFPFLYLIOTITLSSLVKARDNGIMILELINT
FSKQIMGHEYKHIEDFIRTPQWAFLEKLNNAKVRQYIEIYKPKNAVLDKKKK
SSRSKTRVRVQVQVIEIKFNFISSNCVNIENDKKELHRRPIAKCRKEAVSRUKHGQ
ALRKKDENIMAFROEIVDIEIYLTQKQKN"
2246..3106
/notes="IS240"
/complement(2246..2262)
/rpt_type=INVERTED
2337..3044
/genes="pBt003"
2337..3044
/genes="pBt003"
/notes="Similar to Bacillus thuringiensis insertion element
IS240-a protein TR:Q45766 (EMBL:M23740) (235 aa) fasta
scores: E(): 2.2e-92, 99.57% id in 235 aa, and to
Mycobacterium fortuitum, transposase tnp trp6100 TR:Q49185
(EMBL:X53635) (254 aa) fasta scores: E(): 1.4e-37, 48.05%
id in 231 aa"
/codon_start=1
/transl_table=11
/product="insertion element IS240 protein"
/protein_id="CAD330065.1"
/db_xref="GI:21685412"
/db_xref="GOA:Q8KNX4"
/db_xref="InterPro:IPR001584"
/translations="MEKENIFKWKHYQADMILTVRWLYRNLFRDLVEMMBERGSL
LSHTITMRWVHOYGPSELNERIKRHLKRTNDSWRVDETYIKIGENMYLYRAVDSEGGH
LDPLYSKGRDAKAAKCFKLKALASFHVTKPRVITVDGNKAYPPVAIRLKNKESIPYGN
PLRVKKYLNMMIEQDHRFFIKKRIILNMLGLKSMQTAVQVMIAGIEAMHVMKKQLKLRQA
SAQNQRNCITHQLFGLTA"
misc_feature
repeat_unit
repeat_unit
repeat_unit
gene
CDS
/genes="pBt003"
/notes="HMPfam hit to PF00665, Integrase core domain"
3090..3106
/rpt_type=INVERTED
3521..4381
/notes="IS240"
/complement(3521..3535)
/rpt_type=INVERTED
3612..4319
/genes="pBt004"
3612..4319
/genes="pBt004"
/notes="Similar to Bacillus thuringiensis insertion element
IS240-a protein TR:Q45766 (EMBL:M23740) (235 aa) fasta
scores: E(): 3.5e-91, 99.14% id in 235 aa, and to
Mycobacterium fortuitum, transposase tnpA or tnp6100
TR:Q49185 (EMBL:X53635) (254 aa) fasta scores: E():
1.1e-37, 48.05% id in 231 aa"
/codon_start=1
/transl_table=11
/product="insertion element IS240 protein"
/protein_id="CAD330066.1"
/db_xref="GI:21685413"
/db_xref="GOA:Q8KH55"
/db_xref="InterPro:IPR001584"
/translations="MEKENIFKWKHYQADMILTVRWLYRNLFRDLVEMMBERGSL
LSHTITMRWVHOYGPSELNERIKRHLKRTNDSWRVDETYIKIGENMYLYRAVDSEGGH
LDPLYSKGRDAKAAKCFKLKALASFHVTKPRVITVDGNKAYPPVAIRLKNKESIPYGN
PLRVKKYLNMMIEQDHRFFIKKRIIRNMLGLKSMQTAVQVMIAGIEAMHVMKKQLKLRQA
SAQNQRNCITHQLFGLTA"
misc_feature
repeat_unit
gene
CDS
/genes="pBt004"
/notes="HMPfam hit to PF00665, Integrase core domain"
4365..4381
/rpt_type=INVERTED
complement(4634..5275)
/genes="pBt005"
complement(4634..5275)
/genes="pBt005"
/notes="Similar to N-terminus of Bacillus anthracis pxol-18
TR:Q9X2Y9 (EMBL:AF065404) (315 aa) fasta scores: E():
4.2e-56, 84.15% id in 183 aa, and to Bacillus halodurans
Bh234 protein TR:Q9KAC5 (EMBL:AP001515) (378 aa) fasta
scores: E(): 1.6e-18, 35.45% id in 189 aa, and weakly to
Lactobacillus delbrueckii integrase/recombinase orf2
TR:Q48538 (EMBL:Z50864) (333 aa) fasta scores: E(): 6.3,
28.88% id in 90 aa, and to Bacillus thuringiensis
resolvase tnpI SW:TNRI BACTU (P10020) (284 aa) fasta
scores: E(): 8.5, 23.88% id in 180 aa"
/codon_start=1
/transl_table=11
/product="integrase/recombinase family protein"
/protein_id="CAD330067.1"
/db_xref="GI:21685414"
/db_xref="GOA:Q8KNX3"
/db_xref="InterPro:IPR004107"
/translations="MSKNINLSETLAFAFSYLLIEKGRKQSTIKRYAYDIKDFYKWLRA
NKKLHLKSWSELSEADYQTYFSELDKRRISKQTHRIWVVKLHMLFUCIYSPLDGV
INLSLIPDQSLNDNDPTEMBEKLKQTLSTGRQAKYRPLMDRNVCIIVL
NYGLSLQELVSLNMSHINFARNITCYVKKIDTLNRDILFVKQLEIVAFHRTSF"
misc_feature
RBS
gene

```

```

CDS
/genes="pBt006"
complement (5364..5507)
/genes="pBt006"
/notes="Similar to Bacillus anthracis pxol-17 TR:Q9X2Y8
(EMBL:AF065404) (47 aa) fasta scores: E(): 2.1e-12, 68.08%
id in 47 aa"
/codon_start=1
/transl_table=11
/product="putative integral membrane protein"
/protein_id="CAD30068.1"
/db_xref="GI:21685415"
/db_xref="UniProt/TrEMBL:Q8KNX2"
/translations="MAVLKDSNIEMNTYNEIGFFKGFWMGLIFVVPFWFSIMIAITFIW
LCK"
misc_feature
complement (5370..5435)
/genes="pBt006"
/notes="1 probable transmembrane helix predicted for pBt006
by TMHMM2.0"
RBS
complement (5515..5519)
gene
complement (6451..8160)
/genes="pBt007"
CDS
complement (6451..8160)
/genes="pBt007"
/notes="Similar to Bacillus anthracis pxol-16 TR:Q9X2Y7
(EMBL:AF065404) (569 aa) fasta scores: E(): 0, 96.13% id
in 569 aa, and to Bacillus thuringiensis pxol1 orf16-like
protein TR:CAC50562 (EMBL:AJ296638) (310 aa) fasta scores:
E(): 6.5e-122, 99.67% id in 310 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAD30069.1"
/db_xref="GI:21685416"
/db_xref="UniProt/TrEMBL:Q8KNX1"
/translations="MKSFIQNSLNGKITNAVVLTKSWQFVAAPBEKYSYKD
TYLEDILQWTVYERIDVGFVEKAKPSFRRMSTLAETGRKHKLYGITNADLEV
FYLLHKKQTVIPNVITIMMEDYKQYKEEFAYIOHSQFYIALKKLNHNIISIE
GLDGRYIKLTHPMNETEKANPYVVISPVITKDFKLSIAAKKFLDIDAMQOHTET
TLKRSUDKODERNKTHFGKMYRFLHKYVPHQIRAVIEELTALPCTGNPLFKICKMO
KGVHTKRYTLLYLSHSDPLCSKEAGEHQRPFPKATYAKKAFLEAVLQBMNIG
ELSDMNKIHLVLKTHQSVIRIGLRDWDKRGEPYTKIVTTLKLLHQTSQVQI
LDTAKEGILYPLAIIKPERNSDRQAVNFGLYSMYSLHNKIMFRNVHALLKQK
FAPVTEESYNYLYQEBTELPKAYDQGNLHAYIALEIEMREKLKVRGHKERTI
PSDREWFEIADKLPQEKLVTELKQFNLLSFMTFRLVRAGVTITAPDQVLHAM
EIK"
complement (8168..8173)
RBS
8947..8951
RBS
8959..10653
gene
8959..10653
/genes="pBt009"
/genes="pBt009"
/notes="Similar to Bacillus anthracis pxol-14 TR:Q9X2Y5
(EMBL:AF065404) (564 aa) fasta scores: E(): 1.6e-191,
Alignment Scores:
Pred. No.: 5 88e-56 Length: 127923
Score: 897.50 Matches: 247
Percent Similarity: 48.32% Conservative: 126
Best Local Similarity: 31.99% Mismatches: 270
Query Match: 23.20% Indels: 129
DB: Gaps: 30
US-10-783-417-2 (1-735) x BTPBTOXIS (1-127923)
QY 1 MetAsnGlnAsnAsnAsnAsnGluTyrGluIleAspSerHisThrSerProTyr 20
Db ATGAATCCATATCAAAATAAGATGAATATGAATAATTCATGCTCCATCCATGGTTT 39120
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db AGCAAGTCTAATAACTATCTAGATATCCATAGCAAAATAGCCCAACCA 39174
QY 41 LeuGlnAsnThrAsnTyrGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60

```


Db 40117 AGAAATTCGAGACATCTTGTGATCTTCTTATGATATTTTCTTTTACAGTAACAGATG 40176
Qy 403 ValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyr 422
Db 40177 -----GCCTTACACATACTAATAATGATGATCGCAACATAATCTGG 40215
Qy 423 GlyGluArgThrGly-----SerProThrThrIleThrIleArgProPheGluSer 439
Db 40216 GGAGCGGTTTCATGGAATATATTTCTTCACAGACATCCAAAGATATTTCTTTTATAGA 40275
Qy 440 Tyr-----LysValSerIleValThrAspArgGlnSerProProValSerPro 455
Db 40276 AACAAACCTATTGATAAGTCGAATTTGTACAGATAGAGAGTACTCAGATATAATA--- 40332
Qy 456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn 475
Db 40333 -----TATGAAATGATATTTTTCGAATAGCAGTAA 40365
Qy 476 ThrLeuLysTyrSerAlaGlyGlySerLeu---SerAsnTyrGlnAsnThrThrPhePhe 494
Db 40366 GTATTTCGATATTCATCCATTCACATAGAAATATATATAAAGAACTGATCTTAT 40425
Qy 495 GlnPheProArgLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPhe 514
Db 40426 ATGATTCCTCAAAACAAACATCGAAAAAT----- 40452
Qy 515 AsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGly 534
Db 40453 AAAGAAATATGCTCATCTATCATATAAATAAAGCTGATATATATTTTTCAGTAGTT 40512
Qy 535 LeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArg 554
Db 40513 AGAAGAAGAGACAGTT-----GCATTAGTTGGACACATATCTAGTGTTGATTTC 40563
Qy 555 TyrAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeu 574
Db 40564 CAAATACATAGATTAGATAACATCACCACCAATCCAGCTCAAAAGCTTGAAGTA 40623
Qy 575 AspThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeu 594
Db 40624 AGTCTGATTCGAAATTTGTGAAGGCTCTGGTCCACAGGTGGAGACTTGGTAATCTT 40683
Qy 595 GlnSerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPhe 614
Db 40684 AAAGATAGTAGTATTTAGAGTAGATTT---TTAAAAAATGTTTCTCGACATATCAA 40740
Qy 615 IleArgLeuArgTyrAlaThrAsnGlyValGlyAsnThrLeuProAsnIleSerLeuThr 634
Db 40741 GTACGATTCGTTATGCTACTAATGCT-----CCAAAGACAAACAGTATTC 40785
Qy 635 IleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsn 654
Db 40786 TTAACCGGAATAGATACATAAGTGTGAG---CTCCCTAGTACCCTTCCCGCCAAAC 40842
Qy 655 TyrAsn-----AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThrVal 672
Db 40843 CCAATGCTACAGATTAAACATATGCAATTTTGGATATGTAACATTTTCCAAGAACAGTT 40902
Qy 673 -----ThrLeuProLeuAsnArgAsnIleProPheIle 683
Db 40903 CCAATATAAACATTTGAAGGAGAGACACTTTTATTAATGACC-----TTA 40947
Qy 684 PheAsnArgAlaAspValSerAsnSerIleLeuIleAspLysIleGluPheIlePro 703
Db 40948 TATGGTACACCAATCATTCATATATATA---TATATTCACAAATCGAATTTATTTCCA 41004
Qy 704 IleThrSerSerMetHisGlnAsnArgGluGlyGlnLysLeuThrIleGlnThrLys 723
Db 41005 ATCACTCAATCTCTATTAGATTATACAGAGAAGCAAAATATAGAAAAACACAGAAAAA 41064
Qy 724 IleAsnThrPhePheThrAsnHisThrLysThrLeu 735
Db 41065 GTGAATGATTTATTGTTTAAATAAAACAAAGTTCTT 41100

RESULT 12

E00614

LOCUS

E00614 3756 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding a polypeptide having insecticidal activity(BTI
endotoxin).

ACCESSION

E00614

VERSION

E00614.1

KEYWORDS

JP 1986005098-A/1.

SOURCE

Bacillus thuringiensis

ORGANISM

Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
cereus group.

REFERENCE

1 (bases 1 to 3756)

AUTHORS

Aran,E.U. and Toomasu,J.P.

TITLE

INSECTICIDAL POLYPEPTIDE

JOURNAL

Patent: JP 1986005098-A 1 10-JAN-1986;

COMMENT

SHINTOKO CORP

OS Bacillus thuringiensis

PN JP 1986005098-A/1

PD 10-JAN-1986

PF 21-FEB-1985

PR 22-FEB-1984

AR 22-FEB-1984

C07K15/04

A01N63/02

C07H21/04

C12N15/00

C12R1:19;

CC strandedness: Single;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

CC *source: strain=Israelensis ONR-60A;

CC *source: clone=BTIA4-1;

FH Key

Location/Qualifiers

5'UTR <1..925

CDS 926..2971

FT /product='BTI endotoxin'

FT mat_peptide 926..2968

FT /product='BTI endotoxin'

FT 3'UTR 2972..3756.

FEATURES

source

1..3756

/organism='Bacillus thuringiensis'

/mol_type='genomic DNA'

/db_xref='taxon:1428'

ORIGIN

Alignment Scores:

Pred. No.: 1.49e-56

Score: 880.00

Length: 3756

Matches: 245

Percent Similarity: 47.85%

Conservative: 123

Best Local Similarity: 31.86%

Mismatch: 279

Query Match: 22.74%

Indels: 122

Gaps: 31

DB:

US-10-783-417-2 (1-735) x E00614 (1-3756)

Qy 1 MetAsnGlnAsnAspAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20

Db 941 ATGAATCCATATCAAAATGAAGATGAATGAATATTCATCAATGCTCCATCCATGTTT 1000

Qy 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40

Db 1001 -----ACCAAGTCTAATACTATTTCTAGATATCCATTAGCAAAATGAACCAATCAACCA 1054

Qy 41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60

Db 1055 CTGAAAAACACCAATTAACAAAGATTGGCTCAATGTGTGTCAAGATAATCAACAATATGC 1114

Qy 61 AspAsnThrGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80

Db 1115 AATAATGCGGGAATTTTGTAGTTCTGAAACTATTGTTGGAGTTAGTCAGGTATTATT 1174

QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
 DB 1175 GTAGTAGGAAGTATGTAGGA-----GCTTTGTGCGCCCT----- 1210
 QY 101 IleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120
 DB 1211 GTCCTAGCTGCAGGTATATATCTTTTGGACATTTGTCCGATCTTTTGG---CAGGA 1267
 QY 121 GluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
 DB 1268 TCTGACCTCGCAAACTGTTGGCAGGATTTGTTAAACATCGA-----GGAAGG 1315
 QY 141 ProLeuThrGlu-----SerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPhe 157
 DB 1316 CCTATACAGAAATAGATAAAACATAATTAATGTACTTCTTCTATCGTAACACCTATA 1375
 QY 158 ArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeu 177
 DB 1376 AAAAATCAACTGTGATAAATATCAAGAAATTTTCGATAAATGGGAGCCAGCACGT--- 1429
 QY 178 GlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIle 197
 DB 1430 -----ACACACGCTAATGCTAAACAGTACATGATCTCTTACT 1468
 QY 198 ArgPheGluAsnValHisAsn-----AspPheIleArgGluIleProGlyPheGln 214
 DB 1469 ACCTTAGAACCTATAATAGATAAAGATTTAGATATGTTAAATAAATAATGCTAGTATCGA 1528
 QY 215 LeuGluThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeu 234
 DB 1529 ATACCACAA-----CTCCCTGCATATGCACAAATAGTACTTGGCACTTG 1573
 QY 235 AsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSer 254
 DB 1574 AATTATTAAACATCTGCTACTTACTTATCAATATATG-----CTGCAAAATCAA 1624
 QY 255 GlnIleGluProAsnAlaGlyThrSerAspAspTyrTyrLys---LeuLeuLysGluAsn 273
 DB 1625 GGTATAAATCCAAGTACTTCAATTCATCTAATTAATCTATCAGGCTATTTAAACGTA 1684
 QY 274 IleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArg 293
 DB 1685 ATACAGAAATATCTGACTATTGTATACAAACCTCAATCAGGACTTAATGATTAGA 1744
 QY 294 AspGluProAsnMetLysTrpSerIlePheAsnAspTyrArgTyrMetThrIleThr 313
 DB 1745 ACTAATACTAACCAACATGGAATATGATATATCTACCTGTTAGAAATGACTTA 1804
 QY 314 ValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIle 333
 DB 1805 GTGTTAGATCTTATTGCTATTTTCCAAATTAATGACCCAGAAAAATAT----- 1852
 QY 334 GlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrGlu 353
 DB 1853 -----CCAATAGGAGTTAAATCTGAATCTTATCAGAGAAGTTTAT---ACGAAT 1897
 QY 354 IleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyr 373
 DB 1898 GTTAATTCAGATACATT-----AGAACCAATACAGAACTAGAAAAAT 1939
 QY 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
 DB 1940 GGATTAACATAGA---AATCCTACATTATTCTTGGATAAACCAAGGCGGTTTTTACACA 1996
 QY 394 GluAsnThrAsn-----PheGlyAsnArgLeu 402
 DB 1997 AGAAATCTCGAGACATTTCTTGATCTCTTATGATATTTTCTTTTACAGGTAACCATG 2056
 QY 403 ValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyr 422
 DB 2057 GCCTTTACATACTAATGATGAT-----CGCAACATAATCTGGCAGCGGTTCAT 2107
 QY 423 GlyGluArgThrGlySerProThrThrLysThrIle-----ArgProPhe 437

DB 2108 GGAATATATTATTCTCAAGACACATCCAAAGTATTCTCTTTTATAGAAACAACTATT 2167
 QY 438 GluSerTyrLysValSerIleValThrAspArgGlnSerProProValSerProIleGln 457
 DB 2168 GAT-----AAGGTCGAAATGTGCAGATAGAGAGTACTCAGATATATAA----- 2212
 QY 458 ProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsnThrLeu 477
 DB 2213 -----TATGAAATGATATATTTTTCGAATAGCAGTGAAGTATT 2251
 QY 478 LysTyrSerAlaGlyGlySerLeu---SerAsnTyrGlnAsnThrThrPhePheGlnPhe 496
 DB 2252 CGATATTTCATCCCAATTCACAATAAGAAATAATATAAAGAACTGATCTTATATGAT 2311
 QY 497 ProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsn 516
 DB 2312 CCAAAACAAACATCGAAAAAT-----GAAAGAA 2338
 QY 517 TyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGln 536
 DB 2339 TATGGTCATACTCTATCGTATATAAAACTGATAATTATATATATTTTCAGTAGTAGAGAA 2398
 QY 537 LeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsn 556
 DB 2399 AGAAGAGAGAGTT-----GCATTTAGTTGGACACATACTAGTGTGTGATTTCCAAAT 2449
 QY 557 AlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThr 576
 DB 2450 ACATAGATTAGTAAATACATCACCAATCCACCTCTTAAAACTTTAAAGTAGTAGTCT 2509
 QY 577 AsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSer 596
 DB 2510 AATTCAAAAATTTGAAAGGCGCTCGGTACACAAAGTGAATACTTGGTAATTTCTTAAAGT 2569
 QY 597 GlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616
 DB 2570 AGTATGAATTTTAGAGTTAGATTT---TTAAAAAATGTTTCTCAACATATCAAGTACGT 2626
 QY 617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
 DB 2627 ATTCGTTATGCTACTAATGCT-----CCAAAGACAAACAGTAGTATCTTAACC 2671
 QY 637 GlyValIleGlyIleProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn 656
 DB 2672 GGAATAGATACTATAAGTGTGGAG---CTCCCTAGTACCCTTCCCGCAAAACCCAAAT 2728
 QY 657 -----AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThrValThrLeu 674
 DB 2729 GCTACAGATTAAACATATGCAGATTTTGGATATTAACATTTTCAAGAACACAGTTTCA--- 2785
 QY 675 ProLeuAsnArgAsnIle-----ProPheIle-----PheAsnArg 686
 DB 2786 -----AATAAACATTTGAAGGAGAGACACTTTTATTAATGACCTTTATATGCTACA 2839
 QY 687 AlaAspValSerAsnSerIleLeuIleAspLysIleGluPheIleProIleThrSer 706
 DB 2840 CCAATCATTCATATAATA---TATATGACAAAAATGAATTTATTTTCCAATCACTCAA 2896
 QY 707 SerMetHisGlnAsnArgGluLysGlnLeuGluThrIleGlnThrLysIleAsnThr 726
 DB 2897 TCTGTATTAGATTATACAGAGAACCAAAATATAGAAAAAACACAGAAAAATAGTGAATGAT 2956
 QY 727 PhePheThrAsnHisThrLysThrLeu 735
 DB 2957 TTATTTGTTTAAATAAAACAAAGTTCTT 2983

RESULT 13

AF285775 4959 bp DNA linear BCT 16-AUG-2000
 LOCUS
 DEFINITION Bacillus thuringiensis subsp. finitimus parasporal inclusion
 protein Cry (cry) gene, complete cds.
 AF285775
 ACCESSION

```

AF285775.1 GI:9828609
Bacillus thuringiensis serovar finitimus
Bacillus thuringiensis serovar finitimus
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE
1 (bases 1 to 4959)
Moore,G.D. and Debro,L.H.
Cloning and Analysis of a Gene for Exoporiom Enclosed Cry Protein of Bacillus thuringiensis subspecies finitimus
Unpublished
JOURNAL
2 (bases 1 to 4959)
Moore,G.D. and Debro,L.H.
Direct Submission
AUTHORS
TITLE
Submitted (07-JUL-2000) Biology, Jacksonville State University, 700 N. Pelham Road, Jacksonville, AL 36265, USA
JOURNAL
FEATURES
Location/Qualifiers
1..4959
/organism="Bacillus thuringiensis serovar finitimus"
/mol_type="genomic DNA"
/sub_species="finitimus"
/db_xref="taxon:29337"
1128..4514
/gene="cry"
1128..4514
/gene="cry"
/codon_start=1
/transl_table=11
/product="parasporal inclusion protein Cry"
/protein_id="AA600235.1"
/db_xref="GI:9828610"
/translations="MDKHNIKIGQSTEVNSESIFPNPEVDSSTVAIVSAGIVVVGT
ILTFAPASFVNPGVLISFGFLADLPWDPDEPKIWSOPMKHGEDLLNOTISTAVKE
IALAHNGKFDVLTYYERGFANLWRPNSANTARLVQSOFENAHFNFSVMNPOLQPTY
YCTLLSCYTEAANLHNLHQGVQFADOWNADOPHPMLKSSGTYDDELLVTIEKYIN
XTCYTHKGLNHLSKESEKI2WDIAINTYREMTLIVLDLVAITFPFYDIRPPRGEVELL
TREVTYSPLPGNAGPIEIDFSYLEHLITRPGLFTWLSDIELTYESVABGDYLSG
IRESKYTGNQFTMKNINYNTNVNLSKQLITLLPGFIHTLSINRGFTQIAGINKLYS
LIQIVFTFKNDNEYQKFNVSNONQEPOTETNNPDYGGSSQKFKNHSLHPFLIIH
QVEFASYFHSIIPALGWTKHNSVNONLISESVTOIPLVKAYEVTTNSVIRPGFTGCD
LIBLRDKCSIKCKASKKVAYSLSFYAANNATAVSI2VDGSGAGVLLOPTFERKGNNN
FTIQDLNYDXDPQVHTLLVIDELPESEIHHLKREDDEYEGVILLIDLKLEFPIDBNY
TNEMLEKAKVANVLFINATNALKMDYTDIDQVANLVECLISDDLYAKBKIKLHFA
IKPAKQSNRLSPDNFNFNNAENSWTANTGVTIIEGPLYKGRAIQLSAARDNF
PTYLYQKIDESLLKPTRYQLRFQVGSODLEDLVRVGATDIDVMNVPDGLLELSYA
PINPCIEIRLDTTGALDRCKOSVNSAADVRPDVNGDPHAPSFDHDTGTDTNN
RNLGWIIFKLIATPDGYAFGNLELLELGPLSGEALAQQVORKEQWKNTTQKREEA
KLAAAKQIINQLFADQSGTKLRFDFEFNSILDKLVYKIRDVYESVLVSIVPLGNYD
LFMELENRIQNALDLYDARNYTVBFNRGLNANMASSNTEVRQIOQTSVLVLSSWNA
QVAQSLNVRPDPHGCVLRVTAKEBGIGNGVYTLDCANHIDTLTFSSCDSGFTTSNEL
AAVYTKLBIFPDQDIQIRIEIGTEGTFYVESVDLIRMED"
ORIGIN
Alignment Scores:
Pred. No.: 4,46e-55 Length: 4959
Score: 862.50 Matches: 234
Percent Similarity: 50.07% Conservative: 112
Best Local Similarity: 33.86% Mismatches: 252
Query Match: 22.29% Indels: 93
DB: 1 Gaps: 27
US-10-783-417-2 (1-735) x AF285775 (1-4959)
QY 68 SerAlaAspThrIleAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAla 87
Db 1209 AGCAGTGATACGTCGTGAGTGAAGCGCAGGAGATTGTAGTTGGGTACTACTAGCA 1268
QY 88 GlyIleGlyLeuThrSerIleSerGlyProileGlyIleIleGlyAlaIleIle 107
Db 1269 -----GCCTTTGCATCATTTGTTATCCA-----GGTGGGTACTATATA 1307
QY 108 SerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluInAsp----LysThrVal 126

```

QY 479 TyrSerAlaGlySerLeuSerAsnTyr-----GlnAsnThrThrPhePheGln 495
 Db 2349 TATCAAAAAATTTTAAATGTAATATCAAAATGAACCTCAAGAAACTACA----- 2399
 QY 496 PheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsn 515
 Db 2400 -----AACTATCCTTAATGATTATGGTGTCA---AAACGCCAA 2435
 QY 516 AsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeu 535
 Db 2436 AAATTCACACATAATTAATCTCATTTTCCATTAATCATCCAC----- 2477
 QY 536 GlnLeuGlnIleLeuAsp-----ThrGlyValLeuGlyTrpThrHisSer 551
 Db 2478 CAAAGTAGAGTTTGCTGAGTATTTTCACTCTATATTTGCAATAGTTGGACACAATAGT 2537
 QY 552 ValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGly 571
 Db 2538 GTAAACTCCCAAAATTTATATATCAGAAAGTGTGAGTACACAAATCCCATTTGGTAAAGCT 2597
 QY 572 AsnAsnLeuAspThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeu 591
 Db 2598 TACGAGTTACTACCAATTCAGTATATGAGGACGAGTTTACAGTGGAGATTTA 2654
 QY 592 ValTyrLeuGlnSerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGln 611
 Db 2655 ATAGAACTTCGRGAT-----AAATGTTCTATTAATGTAAAGCT---AGTCTCTTAAAA 2705
 QY 612 SerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIle 631
 Db 2706 AAATACGCTATAAGTCTATTTATGCTGCAAAATACGCAATAGCTGTATCAATAGACGA 2765
 QY 632 SerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsnAsnThrPheSer 651
 Db 2766 GGTGATTCGGAGGAGGAGTTCTA-----TTGCAACTACCTTTTCT 2807
 QY 652 GlyThrAsnTyrAsnAsn-----LeuGlnTyrGlyAspPheGlyTyrPhe 666
 Db 2808 AGAAAAAGGAACAATAATTTTACAATTCAGACCTTAACTATAAGGATTTCAATATCAT 2867
 QY 667 GlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArg 686
 Db 2868 ACACCTTTTAGTTGATATTTGAATTTACCCGAAAGTGAAGAAATTCATATCCATTTGAAGCGA 2927
 QY 687 AlaAsp-----ValSerAsnSerIleLeuIleAspLysIleGluPheIleProIle 704
 Db 2928 GAGGATGATTATGAGGAGGAGTGTATTTTAAATGATAAATAGAGTTCAACCTATA 2987
 QY 705 ThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLysIle 724
 Db 2988 GATGAAATATAT-----ACTAATGAATGAATTTAGAGAGGCAAGCAAGCAGTG 3038
 QY 725 AsnThrPhePheThrAsnHisThrLysThrLeu 735
 Db 3039 AATGTATTATTATAAACCAACAAACGCTTTG 3071
 RESULT 14
 BTH251977
 LOCUS BTH251977 6009 bp DNA linear BCT 22-DEC-2001
 DEFINITION Bacillus thuringiensis subsp. medellin cry29Aa gene for Cry29Aa protein.
 ACCESSION AJ251977
 VERSION AJ251977.1 GI:17977978
 KEYWORDS Cry29Aa gene; Cry29Aa protein.
 SOURCE Bacillus thuringiensis serovar medellin
 ORGANISM Bacillus thuringiensis serovar medellin
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
 REFERENCE 1
 AUTHORS Delecluse, A. and Orduz, S.
 TITLE Characterization of two new mosquitocidal toxins, Cry29A and Cry30A, from Bacillus thuringiensis medellin

JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

Unpublished
 2 (bases 1 to 6009)
 Delecluse, A.
 Direct Submission
 Submitted (22-DEC-1999) Delecluse A., Bacteries & Champignons
 Entomopathogenes, Institut Pasteur, 25, rue du Dr Roux, 75724 Paris
 Cedex 15, FRANCE

FEATURES
 source

1. 6009
 /organism="Bacillus thuringiensis serovar medellin"
 /mol_type="genomic DNA"
 /isolates="163-131"
 /sub_species="medellin"
 /db_xref="taxon:79672"
 /country="Colombia"
 642. 2594
 /gene="cry29Aa"
 642. 2594
 /genes="cry29Aa"
 /function="putative mosquitocidal toxin"
 /codon_start=1
 /evidence=experimental
 /transl_table=11
 /product="Cry29Aa protein"
 /protein_id="CAC80985.1"
 /db_xref="GI:17977979"
 /db_xref="GOA:Q8VNX2"
 /db_xref="UniProt/TREMBL:O8VNX2"
 /translation="MNSFQNKVEYLINAPSSISNIPNNYSKYPIANTILNQTOMTNY
 KWTNWCALDNNLKSNPPFANLQNSLIVGLFATAIASLLSAPITGTSIAAGTAIA
 AAIPLPWPQSNENLPDLKLAISEATLYSFQDQREDALTRLESKDSVKVFNFAFT
 FIMNPSNTNTTTFRRFQVNGFVSGMAFFRAKNYEPILLSTYAQAARLHLHLD
 GITYAASKWLSROGDDGLYKFNKCYNEVIEHCIKWNBSLSLLKSGVANNLEY
 NOVYRPLTASVLVDVLSLFSYDPLRYKERSVILTKLYTDPINVHRGISTLEADESK
 YLLEPLTQLYTLTFYSNIFPNYMGHTWYRILSPDKIPAESRFGKSSYIDKVPVI
 PNDKSLIYKIRAYDNHNGLFNVNFGFDGKQDQIKIIGSSSTEIVIKNCKTRLADV
 ISHDLDEKNCYSFAWSTTISLENIKNDITQIPAVKAYQLGVQSQVKPGPHTGG
 NTLNKSNDYLRISCOHLSNVTKYFVIRYATNGSLNTRPIINITIPGTPQGMVLD
 NTFSGTGYNLEQFYKFEFLKEVTLNPNQISLTLNRSQDNQNSILLDLRIEFLPI
 TPSIRKSEQNLEKNQKTVNKLFFN"

ORIGIN

Alignment Scores:

Prod. No.: 5.01e-54 Length: 6009
 Score: 850.00 Matches: 232
 Percent Similarity: 46.97% Conservative: 124
 Best Local Similarity: 30.61% Mismatches: 276
 Query Match: 21.97% Indels: 126
 DB: 1 Gaps: 22

US-10-783-417-2 (1-735) x BTH251977 (1-6009)

QY 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
 Db 642 ATGAATTCATTTTCAAAATAAAAATGAATATGAATAATATAATGCTCCATCGAGT----- 695
 QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
 Db 696 ATTTCTTAATATACCTAATAATATTCTTAAATATCCAAATAGCTAATGATAATAATACTAAACT 755
 QY 41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCys-----GlnGlyAsnThrGln 58
 Db 756 ATGCAAAACACCAATTTACAAAGATTGAGCTAATATGTCGCAATGATAATAATACTAAAA 815
 QY 59 TyrGlyAspAsnPheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGly 78
 Db 816 TCTATAAATCCTTTTGAAGCTAATCTTCAAAATAGTTTAGTAGTCTTTTGTCTATAACT 875
 QY 79 ThrIleValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyPro 98
 Db 876 GCTGCTATTGCTTCATTGTTGTCAGCTCCCAATTAAGTGGTGATCATCTATAGCAGCTGGC 935
 QY 99 IleGlyIleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpPro 118

```

Db      936  ACAGCTATAGCAGCAGCTATA-----ATACCTTACTTTGGCCT 974
Qy      119  AlaGlyGluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheVal 138
Db      975  TCT-----CAAGAAATAATACTTCCCGATAAATCTTAGCTATTTCAGAAACCAACCTT 1028
Qy      139  AspThrProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArg 158
Db      1029  TATTCATTTCAAGATCAAGGGTTAGAGAGATGCACCTTACGCCGATTAGAAAGTTTAAAA 1088
Qy      159  GlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGln 178
Db      1089  GATAGTGTAAATATTTTGGAAATGCCTTCACATTTTGGATAAT----- 1133
Qy      179  AlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArg 198
Db      1134  -----AATCCTTAATTTCAACC-----AACCAACTACTGTGCAGAGAAAGA 1172
Qy      199  PheGluAsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyr 218
Db      1173  TTCACAGAGTTTAAATGTAGATTGTAGGCAGCATGGCCTTCTTTAGACAAAGAACTAT 1232
Qy      219  LysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGln 238
Db      1233  GAACCAATACGTGTATCACTTATGCACAGCTGCAAGGTGTACATTTACTTCATTTCGG 1292
Qy      239  GlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluPro 258
Db      1293  GATGGAATTACGTATGCTGAAAAATGGAATTTAAGTAGACAAGGGGATGCATG----- 1346
Qy      259  AsnAlaGlyThrSerAspAspTyrTyrLysLeuLysGluAsnIleProLysTyrSer 278
Db      1347  -----CCCGGGGATTTACTTTTACAAAGAAATTTAACAAGATTTTGAATGATATATT 1397
Qy      279  AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMet 298
Db      1398  GACATTGTATAAGTGGTCAATGAAGTTTGCTCTTCTTAA-----AGTGTAGGTGCA 1454
Qy      299  LysTrpSerIlePheAsnAspTyrArgTyrMetThrIleThrValLeuAspThrIle 318
Db      1455  AATTTGGCTAGATAACAATCAATATCGTACATTTCTTAACAGCCTCTGTATTAGATGTCATT 1514
Qy      319  SerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyIleGluVal 338
Db      1515  TCTCTTTTTTCATCATATGACCCACGTTTATATAAAGAAAGATTA-----AGTGTGAGATT 1571
Qy      339  LysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArg 358
Db      1572  -----CTCAAGAAAAATTTATATACAGATCCAAATTAATTTATCATAGA 1613
Qy      359  LeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAla 378
Db      1614  GGCATTTTCACTAGACAGATGAAGCAATATATACCTTGAGCAACTTTGTTTACTCAA 1673
Qy      379  SerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsnPhe 398
Db      1674  CTATATACACTAATCAATTTTATAGTAATATATTTTATAATTTACATGGGACACACCAAT 1730
Qy      399  GlyAsnArgLeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThr 418
Db      1731  -----ACTTATAGATATTTATCACCG 1751
Qy      419  GluThrLeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGlu 438
Db      1752  GACAAAATTTTTCAGAGCGCTCTTTTGGAAAAACAAGACGCTATATAGAT----- 1802
Qy      439  SerTyrLysValSerIleValThrAspArgGlnSerProValSerProIleGlnPro 458
Db      1803  -----AAAGTCCCTGTAATTCCTACGATAATCA----- 1832
Qy      459  HisPheIleLeuAsnGlnIleGluLeuTyrLeuAsn----- 470
```

```

Db      1833  -----ATTATTTATAAAATTCGACCATATGATAATCATAAATGGTTTATTATTAATGTAATG 1886
Qy      471  -----GlySerSerAsnAsnThrLeuLysTyrSerAlaGlyGlySer 484
Db      1887  TACTTTGGTGTTTTGGGACGAGAAAAAGATCAATACAAAAATAATTAATTCGAGGTAGTTCT 1946
Qy      485  LeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArgLysLysAspCysAsnLeu 504
Db      1947  ACAGAAATATAC-----ATTAATAATTTGTACT--- 1973
Qy      505  ValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSerHis--- 523
Db      1974  -----CATCGCTTAGCAGATGTCATCTCACAATGAT 2003
Qy      524  -----PheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuIleLeuAsp 541
Db      2004  CTTGATGAAAAAACAATAATGTTATTTCTTT----- 2033
Qy      542  ThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerAspLys 561
Db      2034  -----GCATGGACATCTACTCAATTTCTTCTTGAATAATCAAAATAAAAAATGAT 2081
Qy      562  IleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIle 581
Db      2082  ATTATCACACAAATCCCTGCTGTAAAGCCTATCAATTAGGTGTGCAATCACAGTGATT 2141
Qy      582  GluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGlu 601
Db      2142  AAGGTCAGGTCTCATCTACTGTTGGAGACTTAATAATTTTAAATCAAAATGATTATTACGT 2201
Qy      602  IleThrCysGlu-----ThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAla 620
Db      2202  ATATCATGTCAACACTTATCTAATGTTTACAAAAAATAATTTTGTAGAATTCGTACGCT 2261
Qy      621  ThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGly 640
Db      2262  ACAATGTTTCTCTAAATACTAGGCCCAATTAATAATTTACAAATTTCCA-----GGA 2312
Qy      641  IleProGlnArg-----LeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeu 658
Db      2313  ATGACACCACCAAGGTATGGTACTCGATAATACTTTTCTGGTACAGTTTATAGCAATTTA 2372
Qy      659  GlnTyrGlyAspPheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArg 678
Db      2373  GAATATCAAAATTTTGGCTATAAAGAAATTTCTAAAGAGGTAAACATTAATCAAAATCAA 2432
Qy      679  AsnIleProPheIlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleAsp 697
Db      2433  AGTATATCTCTTACATTAATTCGTTCTGTGATCAAAATTCAAATTTCAATTTTACTCTCTGAT 2492
Qy      698  LysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeu 717
Db      2493  AGAATTCGATTTCTACCAATTTACTCCCTCTATACGAAAAAGTAAAGAACACAAATTTTA 2552
Qy      718  GluThrIleGlnThrLysIleAsnThrPhePheThrAsnHisThrLysThrLeu 735
Db      2553  GAAAAAATCAGAAAAACGGTAAACAAAGTTATTTTTCATTAATAATAAAGGGTTG 2606

RESULT 15
AF132928
LOCUS
DEFINITION
  Bacillus thuringiensis subsp. finitimus Cry28Aa1 delta-endotoxin
  gene, complete cds.
ACCESSION
  AF132928
VERSION
  AF132928.1 GI:4574729
KEYWORDS
  Bacillus thuringiensis serovar finitimus
  Bacillus thuringiensis serovar finitimus
  Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
  cereus group.
  1 (bases 1 to 4896)
REFERENCE
  Wojciechowska,J.A., Lewitin,E., Revina,L.P., Zalunin,I.A. and
  Chestukhina,G.G.
```

TITLE Two novel delta-endotoxin gene families cry26 and cry28 from
 JOURNAL *Bacillus thuringiensis* spp. *finitimus*
 PUBMED FEBS Lett. 453 (1-2), 46-48 (1999)
 REFERENCE 10403372
 2 (bases 1 to 4896)
 AUTHORS Wojciechowska, J.A., Lewitin, E.I. and Chestukhina, G.G.
 TITLE Direct Submission
 JOURNAL Submitted (04-MAR-1999) Laboratory of Protein Chemistry, Institute
 of Microbial Genetics, Dorozhny proezd, Moscow 113545, Russia
 Location/Qualifiers

FEATURES
 source 1. 4896
 /organism="Bacillus thuringiensis serovar *finitimus*"
 /mol_type="genomic DNA"
 /strain="VKPM B-1161"
 /sub_species="finitimus"
 /db_xref="taxon:29337"

CDS
 1125..4458
 /codon_start=1
 /transl_table=11
 /product="Cry28Aa1 delta-endotoxin"
 /protein_id="AAD24189.1"
 /db_xref="GI:4574730"

/translation="MAQYKIGVOSTEVNSSEIFNEPVDSDTVAVUSAGIVVVG
 ILTAFAFVNPVGVLIISFGLAPVLPDPEDPKKIWSQFMKGHDLNQTISTAVKE
 IALAHNGFKDVLITYERAFNDWKRNPSANTARLVSRPENAHFNFVSNMPLQLPY
 DTLSCYTEANLNLHGHQVADQWADPHSPMLKSSGTYYDELLVYIEKYIN
 YCTKYHKLHKESEKITWDAYNTYREMTLIVLDLVATPPFDYIRRPGRVLEL
 TRENTSLDLHTRPPLFTWLSIDIELYTESVAEGDYLGLRESKYVTGNQPTMKNIY
 GNTNRLSKQLITLPGEFWTHLSINRPFQTAGINKLYSLIOKIIVPTFPKNDNEVQKN
 FVNNQNPQETITPNDYSGNSQKPKNLHSHFPLIIHKUEFAYFHSIFALGWTHN
 SVNSQLISVESVTOPIKAYEVTVNSVIRGPTGGDLIELDKCSIKKASLKK
 YALISFYAANAIVASIDVDSGAGVLOPTFSRGNNTFTQDLNLYKDFQVHTLLD
 IELPSEIHLKREDDYEGVILLIDKLEFKIDENVTNEMLEKAKAVNVLFIN
 ATNALKMDVTDHIDQVNLVFCISDILYAKEKIKFTPCIKEAKOLSOARNLSPNF
 NNLNENSWANTGVITIEGLDYKRALQLSAARDENPPTLYOKIDESLLKPYTRY
 QLRGVEGSDLDLQVADTIVNVPGDILISYSAPINPCBEIETRLTDTCCAL
 DRCKQSVNSAADYVPQVNGDPHAFSDITGTDYDNNRNLGIWIIKFIATPDGAT
 FGNLELIELGSLGALAQVKEQKQWGNNTQKEEAAKLYAAAKITQINLFADSQG
 TKLRDPTFNSLADLYKIRDVYSEVLVPIGLNYDLFWLELNRIQNALDLYDAR
 NTVTNGEPRNGLANWASNTVEROIQAHPCWYSLGNWAQVAGSLNVKPDHGYLVRT
 AKESIGNGYVITLSCANHIDLTFTSSCDSGFTTSSNELAAVYTKLEIFPPTDQIRI
 EIGETRSTFYVESVDLIRMED"

ORIGIN

Alignment Scores:
 Pred. No.: 4,628-54 Length: 4896
 Score: 849.00 Matches: 229
 Percent Similarity: 48.76% Conservative: 106
 Best Local Similarity: 33.33% Mismatches: 248
 Query Match: 21.94% Indels: 104
 DB: 1 Gaps: 25

US-10-783-417-2 (1-735) x AF132928 (1-4896)

QY 68 SerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAla 87
 DB 1210 ACAGTGATACAGTCGCTAGTAGAAGCGAGGAGTTGTAGTGTGGTACTACTACTGACA 1269
 QY 88 GlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIle 107
 DB 1270 -----GCCTTTGCATCAATTGTTAATCCA-----GGTGTGGTACTTATA 1308
 QY 108 SerPheGlyThrLeuIleThrValPheThrProAlaGlyGluClnAsp---LysThrVal 126
 DB 1309 TCATTGGAACTTGGCTCCCGTCTCTTTGGCTGTATCCAGAGGAAGATCAAAAAAAT 1368
 QY 127 TrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeuThrGluSerIle 146
 DB 1369 TGGTCACATTTATGAAACACGAGAGACCTTTTAAATCAACAATTTCTACAGCTGTA 1428
 QY 147 LysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGlnSerThrAsnThr 166
 DB 1429 AAGAAATAGCATAGCTCATCTAAATGTTTAAAGATGTATTAACTACTACTATGAAGA 1488

QY 167 AlaLeuAspAspThrArgLysLeuLysArgLeuGlnAlaProGlyLeuProSerSer 186
 DB 1489 GCATTAATGATGGAAGAGA-----AATCCAAGTGCA 1521
 QY 187 AlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnApphe 206
 DB 1522 -----AATAGTCCAGATGGGTATCACAGAGATTTGAAACCGCTCATTTCAATTTT 1572
 QY 207 IleArgGluIleProGlyPheGlnLeuGluThrThrLysThrLeuLeuProIleTyr 226
 DB 1573 GTAAGCAATATGCCCAACTCCAACTTCCACGTATGACACATATTATTAAAGTGTAT 1632
 QY 227 AlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGlu 246
 DB 1633 ACAGAAAGCTGCAAAATTTACATTTGAATTTATTACATCAAGGTGTACAATTCGGCGATCA 1692
 QY 247 TrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTyr 266
 DB 1693 TCGAATGCGAGATCAACACCATTTCCCAATTTGTAAGATCATCAGGTACT-----TAT 1743
 QY 267 TyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyrArg 286
 DB 1744 TATGACGACTATTGGTATATTATGAAAGTATATTAAATTATGACCAAGACATACCAT 1803
 QY 287 ThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTyr 306
 DB 1804 AAAGGATTGAATCACCTTAAAGAAATCAGAAAAATCACATGGGATGCTTATATAACATAT 1863
 QY 307 ArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIle 326
 DB 1864 CGTCGAGAAATGACCTTAATTTGGTATCTTGTGCGAACTTTCTTTTATGATATA 1923
 QY 327 LysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThr 346
 DB 1924 CCGTGGTTT-----CCAGGAGGAGTAGAAGCTAGAAATTAACA 1959
 QY 347 ArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeuArgValGlnPro 366
 DB 1960 ACAGAGGTTTATACAAAGTTTAGATCATTTAAACACGACCA----- 2001
 QY 367 AsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeu 386
 DB 2002 -----GGGCTATTACTTGCTG 2019
 QY 387 GluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIleSer 406
 DB 2020 TCAGATATTGAGTTATACAGGAGGTGTGGCAGAGGCGAATTTATTCAGGTATT--- 2076
 QY 407 AsnArgAspAlaProThrTyrSerAsnThr-----IleThrGluThrLeuTyrGly 423
 DB 2077 ---CGAGAGCTAAATATTATCTGTAATCAATTTTACGATGAAAAATATTATGCT 2133
 QY 424 GluArgThrGly---SerProThrThrLysThrIleArgProPheGluSerTyrLysVal 442
 DB 2134 AATACAAATAGATTAGTAAGCAGCTCATTTACATTTATCCAGCGCAA---TTTATGACT 2190
 QY 443 SerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIleIle 462
 DB 2191 CACTTAAGCATAAACCGTCTCTTTTCAACATAGCTGGTATAAATAGTTATACAGTTTA 2250
 QY 463 AsnGlnIleGluLeuLeuAsnGlySerSerAsnAsnThrLeuLysTyrSerAlaGly 482
 DB 2251 ATTCAAAAATCGTATTACAACTTTTAAACAGATAAT-----GAATATCAAAAAAT 2304
 QY 483 GlySerLeuSerAsnTyr-----GlnAsnThrThrPhePheGlnPheProArgLys 499
 DB 2305 TTTAATGTGAATAATCAAAATGAACCTCAAGAAATACATA----- 2343
 QY 500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHis 519
 DB 2344 -----AATATCTTAATGATTATGGTGTTC-----AACGCCAAAAAATTTCAACAT 2391

Qy	520	IleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuIle	539
Db	2392	AATTATCTCATTTTCCATTAAATCATCCAC	2433
Qy	540	LeuAsp-----ThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyr	555
Db	2434	GCTGAGTATTTTCACTCTATATTTGCAATTAGTGTGGACACACATAGTGTAAACTCCCAA	2493
Qy	556	AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleIleGlyAsnAsnLeuAsp	575
Db	2494	AAITTAATATACAGAAGTGTGACGTACACAAATCCATTTGGTAAAGCTTACGAAGTT---	2550
Qy	576	ThrAsnSerLysValIleGluGlyProGlyHisThrGlyLeuValTyrLeuGln	595
Db	2551	ACTAACCAATTCAGTTATTAAGAGGACCAAGCTTTTACAGTGCAGATTTAAATAGAACTTCGT	2610
Qy	596	SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIle	615
Db	2611	GAT-----AAATGCTTCTATTAAATGTAAAGCT---AGTTCTTTTAAAAAATAACGCTATA	2661
Qy	616	ArgLeuArgTyrAlaThrAsnGlyValGlyAsnThrLeuProAsnIleSerLeuThrIle	635
Db	2662	AGTCTATTATTGCTGCAATTAACGCAATAGCTGTATCAATAGACGTAGGTGATTCGGA	2721
Qy	636	ProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyr	655
Db	2722	GCAGGAGTTCTA-----TTGCAACTACCTTTTCTAGAAAAGGGAAC	2763
Qy	656	AsnAsn-----LeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSer	670
Db	2764	AATAAATTTTACAATTTCAAGACCTTAACATAAAGGATTTTCAATATCATACACATTTTAGTT	2823
Qy	671	ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAsp	688
Db	2824	GATATGCAATTTACCCGAAAGTGAAGAATTCATATCCATTTTGAAGCGAGAGATGATTAT	2883
Qy	689	ValSerAsnSerIleLeuIleIleAspLysIleGluPheIleProIleThrSerSerMet	708
Db	2884	GAGGAGGAGTGATCTTTTAATTTGATTAATTTAGAGTTCAACCTATAGATGAATAATTAT	2943
Qy	709	HisGlnAsnArgGluThrLysLysLeuThrIleGlnThrLysIleAsnThrPhePhe	728
Db	2944	-----ACTAATGAAATGAAATTTAGAGAAGCAAGAAAGACAGTGAATGATTATTTT	2994
Qy	729	ThrAsnHisThrLysThrLeu	735
Db	2995	ATAAACGCAACAAACGCTTTG	3015

Search completed: December 12, 2005, 02:22:23
Job time : 9851 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 11, 2005, 20:13:22 ; Search time 1089 Seconds

(without alignments)
4498.209 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869

Sequence: 1 MNQNNNDNEVEIIDSHTSPY.....KLETIQTKINTFFNHTKTL 735

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10783417/runat_01122005_141956_24301/app_query.fasta_1.903
-DB=N Geneseq -QFMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40 cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10783417 @cgn 1.1 752 @runat_01122005_141956_24301 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 21.*

1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3869	100.0	2208	13	ADR89394
2	3314	85.7	2235	13	ADR89396
3	3104.5	80.2	2085	13	ADR89398
4	1061.5	27.4	4571	1	AAN93059

5	1060.5	27.4	3543	2	AAQ14669
6	1060.5	27.4	3543	2	AAQ81178
7	1059.5	27.4	3940	1	AAN93054
8	1059.5	27.4	4934	1	AAN81490
9	1032.5	26.7	2130	12	ADP71294
10	1032.5	26.7	2246	12	ADP71295
11	899.5	23.2	2061	2	AAQ14670
12	899.5	23.2	2061	2	AAQ81180
13	881	22.8	3756	1	AAN50525
14	849	21.9	4896	6	AAD43974
15	849	21.9	4896	10	ADF31301
16	849	21.9	4896	10	ADF31306
17	775.5	20.0	3507	2	AAT43221
18	759	19.6	3684	6	ABK14949
19	759	19.6	3684	14	ABK61392
20	758	19.6	3684	1	AAN82106
21	757.5	19.6	3535	1	AAN91003
22	747.5	19.3	4184	1	AAN96139
23	711.5	18.4	4065	1	AAN92515
24	705	18.2	4391	14	AEA61394
25	703	18.2	3633	6	ABK87235
26	703	18.2	3633	12	ADL15306
27	703	18.2	3633	14	ABK45608
28	703	18.2	6613	6	ABK87248
29	703	18.2	6613	12	ADL15321
30	703	18.2	6613	14	ABK45623
31	696.5	18.0	10974	1	AAN60055
32	689.5	17.8	4188	8	ABX13487
33	689.5	17.8	4188	14	ABK90816
34	678.5	17.5	2003	6	ABK87237
35	678.5	17.5	2022	6	ABK87242
36	678.5	17.5	2022	12	ADL15316
37	678.5	17.5	2022	14	ABK45618
38	677	17.5	2478	12	ADP43079
39	676	17.5	1953	3	AAQ43223
40	676	17.5	1953	3	AAQ09499
41	676	17.5	2425	2	AAQ05679
42	675	17.4	3621	6	ABK87234
43	675	17.4	3621	12	ADL15304
44	675	17.4	3621	14	ABK45606
45	675	17.4	3621	14	ABK34684

ALIGNMENTS

RESULT 1	
ADR89394	
ID	ADR89394 standard; cDNA; 2208 BP.
XX	
AC	ADR89394;
XX	
DT	18-NOV-2004 (first entry)
XX	
DE	AXMI-006 coding sequence.
XX	
ss	ss; gene; delta-endotoxin; delta-endotoxin associate polypeptide;
KW	expression cassette; transformation; transgenic; plant; bacteria;
KW	lepidoptera; coleoptera; pest; pesticide; resistance;
KW	pesticidal activity.
OS	Bacillus thuringiensis.
XX	
FH	Key Location/Qualifiers
FT	CDS 1..2208
FT	/*tag= a
FT	/product= "AXMI-006"
XX	
PN	WO2004074462-A2.
XX	
PD	02-SEP-2004.
XX	
PP	20-FEB-2004; 2004WO-US005829.
XX	

PR 20-FEB-2003; 2003US-0448632P.
PR 20-FEB-2003; 2003US-0448633P.
PR 20-FEB-2003; 2003US-0448797P.
PR 20-FEB-2003; 2003US-0448806P.
PR 20-FEB-2003; 2003US-0448810P.
PR 20-FEB-2003; 2003US-0448812P.
PR 19-FEB-2004; 2003US-00781979.
PR 19-FEB-2004; 2004US-00782020.
PR 19-FEB-2004; 2004US-00782096.
PR 19-FEB-2004; 2004US-00782141.
PR 19-FEB-2004; 2004US-00782570.
PR 19-FEB-2004; 2004US-00783417.

XX (ATHE-) ATHENIX CORP.

PA Carozzi N, Hargiss T, Koziel MG, Duck NB, Carr B;

XX PI

XX WPI; 2004-635574/61.

DR P-PSDB; ADR89395.

XX New isolated delta-endotoxin and delta-endotoxin-associated nucleic acids

PT and polypeptides, useful for killing lepidopteran or coleopteran pests or

PT for producing organisms with pesticide resistance.

XX Claim 1; SEQ ID NO 6; 178pp; English.

XX This sequence encodes an isolated delta-endotoxin. Some of the delta-
CC endotoxin coding sequences of the invention have alternative start
CC codons, producing more than one protein from a single open reading frame.
CC The nucleic acid sequences of the invention are useful in DNA constructs
CC or expression cassettes for transformation and expression in plants and
CC bacteria. The nucleic acids and corresponding polypeptides are useful for
CC killing lepidopteran or coleopteran pests. Compositions containing the
CC delta-endotoxins of the invention, and methods for their production, are
CC useful for the production of organisms with pesticide resistance,
CC specifically bacteria and plants. These organisms are useful for
CC generating altered or improved delta-endotoxin or delta-endotoxin-
CC associated proteins that have pesticidal activity, or for detecting the
CC presence of delta-endotoxin or delta-endotoxin-associated proteins or
CC nucleic acids in products or organisms.

XX SQ Sequence 2208 BP; 842 A; 377 C; 308 G; 681 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 2208
Score: 3869.00 Matches: 735
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-783-417-2 (1-735) x ADR89394 (1-2208)

QY 1 MetAsnGlnAsnAspAsnAsnGluTyrGluIleAspSerHisThrSerProTyr 20
Db 1 ATGAATCAAAATAACGATAATAACGAATATGAAATATTGATTCGCATACCTCACCTTAT 60
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 61 TTTCGAAACAGAAACAGTAAATGATTCAGATACCTTACACAAATAATCCAAATCAACCA 120
QY 41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
Db 121 TTACAAACACAAATTACAAAGAGTGCCTCAATATGTCAGGGAATACACAAATATGGT 180
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80
Db 181 GATAATTCGAGACATTTGCTAGTCTGATACAAATGCTCAGTTAGTCAGGTACTATT 240
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
Db 241 GTATCCGGTACTCTGTAGCCGGTATAGGTGGGCTCACTTCTATATCCGACCGATAGGA 300

QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGly 120
Db 301 ATAATAGGTCTATAAATAATATCTTTTGGTACCCCTAATCACTGCTTTTGGCCCGGGA 360
QY 121 GluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db 361 GAACAGACAAACAGTATGGACAAATTTATAAATGGAGAAATTTTGTGTATACA 420
QY 141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
Db 421 CCGTTAACAGAAAGCATAAACAGTAAAGTTTACAACTTTTGAAGGATTTAGACAAATA 480
QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaPro 180
Db 481 TTACAAAGCTATATAACAGCATTAGATGATGGAGAAATTTAAAGAGACTCAAGCTCCT 540
QY 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuIleValArgPheGlu 200
Db 541 GGATTACCCACCATCATCAGCATTACAAAGCTGCTTGAAGCTTTAAATACGATTTGAG 600
QY 201 AsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThr 220
Db 601 AATGTTCACTAATGATTTTATTCGGAATATACCTGGTTTCCAACTTGAAGCTTATAA 660
QY 221 LeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGly 240
Db 661 CTATTACTACTATTTATGCGCAAGCTGCTAATTTTCAATTTAAATTTATTTACACAGGT 720
QY 241 AlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAla 260
Db 721 GCTGAATGCTGATGAATGGAATGCAGATATATACCTTTCACAAATTTGAACCTTAAT 780
QY 261 GlyThrSerAspAspTyrTyrLysLeuLysGluAsnIleProLysTyrSerAsnTyr 280
Db 781 GGAACATCATGATGATTTATTAATTTTAAAGAAATATACCTAATAATATAGTAACAT 840
QY 281 CysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTyr 300
Db 841 TGTGCAAAATACCTATAGAACAGGACTAAAAAATCTTAGAGACGAAACCAAAATATGAA 900
QY 301 SerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGln 320
Db 901 AGTATATTTTAAATGACTATCGAAGATATATGACCAATTTGTTATAGTACCATCTCT 960
QY 321 PheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluValLysGly 340
Db 961 TTTTCTTTATGATATAAAGATATAGAGATTCATAGGAGGATAGAGTAAAGTAAAGGC 1020
QY 341 IleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuPro 360
Db 1021 ATTAAGAATGAACCTCACAGAGAAATTTATACAACTGAAATAAATTTTGTGCTTCTCT 1080
QY 361 GlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPhe 380
Db 1081 CAACCTTAGAGTTTCAACCCCAATCTAGCTACGATGGAATATATAATTTAACACGTCGA 1140
QY 381 LysLeuPheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsn 400
Db 1141 AAATATTTTCAATTTTATAGAACAAATTTATTTTATACAGAAATATACAAATTTCCG 1200
QY 401 ArgLeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
Db 1201 CGTTTAGTGTGTTTCTTAATCGTGATGCACCTACTTATAGCAATACTATAACTGAACT 1260
QY 421 LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440
Db 1261 TTATATCGAGAAAGAACAGGTTTCAACCCCAACCAATTAAGACCACTTTGAATCTTAT 1320
QY 441 LysValSerIleValThrAsnArgGlnSerProProValSerProIleGlnProHisPhe 460
Db 1321 AAAGTTTCAATTTGATGATAGACAAATCACTCTGTTTCCCTATTTCAACACACCTTT 1380
QY 461 IleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsnThrLeuLysTyrSer 480

```
Db 1381 ATAATAATCAAAATTTGAACCTTTATTAATGGCTCATCTAACACACATCAATATTC 1440
Qy 481 AlaGlySerLeuSerAsnThrGlnAsnThrPhePheGlnPheProArgLysLys 500
Db 1441 GCAGGAGGGCTTTATCTAATATTAATCAAAACACAACTTTTTCATTTCTAGAAAAA 1500
Qy 501 AspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnThrSerHis 520
Db 1501 GACTGCATCTAGTTATGATCCAGGTGTTCCAAACCTTTAATACTATAGTCATATT 1560
Qy 521 LeuSerHisPheSerLeuPheThrThrSerTyrValIleGlyLeuGlnLeuIleLeu 540
Db 1561 TTATCCATTTTTCATTTATTTACTTATCTCTATGTGATTCAGTTACAGCTACAAATATTA 1620
Qy 541 AspThrGlyValLeuGlyTyrPheHisSerSerValAspArgTyrAsnAlaIleSerAsp 560
Db 1621 GATACAGGTGATTAGGATGCACACAGTAGTGTGATAGATATATGCAATATCAGAT 1680
Qy 561 LysIleIleThrMetIleProAlaIleIleGlyAsnAsnLeuAspThrAsnSerLysVal 580
Db 1681 AAAATAATTAACAATGATCCAGCAATCAAGGTAAACAATCTTGATACAAACTCTAAGGTA 1740
Qy 581 IleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeu 600
Db 1741 ATTGAAGGACCTGGTCTATACAGGAGGAACCTTGTTTATTACAAAGTCAAGGCGTTTA 1800
Qy 601 GluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAla 620
Db 1801 GAAATTCATGTCGAACTCTTAATCTTACACATCTTATTTTCATTAGCTCGATATGCT 1860
Qy 621 ThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGly 640
Db 1861 ACAATGGTCTCGAAATACCTCTTCCTAATATATCTCTTACAATACCAGGAGTAATAGGA 1920
Qy 641 IleProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyr 660
Db 1921 ATACCACCTCAAGACCTCAACACACACTTTTCTGGTACAAATTAATAATTTACAAATAC 1980
Qy 661 GlyAspPheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIle 680
Db 1981 GGAGATTTGGGTATTTCCATTTCCAGTACAGTACATACCTTTTAATCGAAACATA 2040
Qy 681 ProPheIlePheAsnArgAlaAspValSerAsnSerIleLeuIleAspLysIleGlu 700
Db 2041 CCATTTATATTTAATCGTCAGATGTATCAAAATTCAAATTTTAATCAATGATAAAATGAA 2100
Qy 701 PheIleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIle 720
Db 2101 TTTATACCAATTTACTTCTCTATGCACCAAAATAGAGAAAAACAAATTTAGAACTATC 2160
Qy 721 GlnThrLysIleAsnThrPhePheThrAsnHisThrLysThrLeu 735
Db 2161 CAAACAAAAATAAATACATTTTTCACAAATCATACAAAAACATTT 2205
```

RESULT 2

ADR89396

ID ADR89396 standard; cDNA; 2235 BP.

XX AC

XX ADR89396;

XX DB:

18-NOV-2004 (first entry)

XX DE

XX AXMI-007 coding sequence.

XX KW

ss; gene; delta-endotoxin; delta-endotoxin associate polypeptide;

KW expression cassette; transformation; transgenic; plant; bacteria;

KW lepidoptera; coleoptera; pest; pesticide; resistance;

KW pesticidal activity.

XX OS

Bacillus thuringiensis.

XX FH

Key Location/Qualifiers

```
FT CDS 1..2235
FT /tag= a
FT /product= "AXMI-007"
FT /transl_except= pos:1..3, aa:Met
FN WO2004074462-A2.
XX
PD 02-SEP-2004.
XX
PF 20-FEB-2004; 2004WO-US005829.
XX
PR 20-FEB-2003; 2003US-0448632P.
PR 20-FEB-2003; 2003US-0448633P.
PR 20-FEB-2003; 2003US-0448797P.
PR 20-FEB-2003; 2003US-0448806P.
PR 20-FEB-2003; 2003US-0448810P.
PR 20-FEB-2003; 2003US-0448812P.
PR 19-FEB-2004; 2004US-00781979.
PR 19-FEB-2004; 2004US-00782020.
PR 19-FEB-2004; 2004US-00782096.
PR 19-FEB-2004; 2004US-00782141.
PR 19-FEB-2004; 2004US-00782570.
PR 19-FEB-2004; 2004US-00783417.
XX (ATHE-) ATHENIX CORP.
XX
PI Carozzi N, Hargiss T, Kozziel MG, Duck NB, Carr B;
XX WPI: 2004-635574/61.
XX P-PSDB; ADR89397.
XX
FT New isolated delta-endotoxin and delta-endotoxin-associated nucleic acids
FT and polypeptides, useful for killing lepidopteran or coleopteran pests or
FT for producing organisms with pesticide resistance.
XX
PS Claim 1; SEQ ID NO 8; 178pp; English.
XX
CC This sequence encodes an isolated delta-endotoxin. Some of the delta-
CC endotoxin coding sequences of the invention have alternative start
CC codons, producing more than one protein from a single open reading frame.
CC The nucleic acid sequences of the invention are useful in DNA constructs
CC or expression cassettes for transformation and expression in plants and
CC bacteria. The nucleic acids and corresponding polypeptides are useful for
CC killing lepidopteran or coleopteran pests. Compositions containing the
CC delta-endotoxins of the invention, and methods for their production, are
CC useful for the production of organisms with pesticide resistance,
CC specifically bacteria and plants. These organisms are useful for
CC generating altered or improved delta-endotoxin or delta-endotoxin-
CC associated proteins that have pesticidal activity, or for detecting the
CC presence of delta-endotoxin or delta-endotoxin-associated proteins or
CC nucleic acids in products or organisms.
XX
SQ Sequence 2235 BP; 861 A; 364 C; 316 G; 594 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 9.25e-290 Length: 2235
Score: 3314.00 Matches: 638
Percent Similarity: 91.28% Conservative: 32
Best Local Similarity: 86.92% Mismatches: 60
Query Match: 85.66% Indels: 4
DB: 13 Gaps: 3

US-10-783-417-2 (1-735) x ADR89396 (1-2235)
Qy 1 MethAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 1 GTGAATCAAAATAAT---AATAATGAATATGAGATTATCGATTCAAGAAATTTATCTTAT 57
Qy 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 58 CTTTCTAACAGAAATATTGTCATCTTAGATACCTTACACAAATAATCAAAATCAACCA 117
Qy 41 LeuGlnAsnThrAsnTyrLysGluThrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
```

118 TTACAAAACAAATTTACAAAGAGTGGCTCAATATATGTCTCAAGGAATACACAAATATGGT 177
QY
61 AspAsnPhelThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80
Db
178 GATAAATTCGAGACATTTGCTAGTGTCTGATACAAATGCTGCAGTTAGTCAGGTACTATT 237
QY
81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
Db
238 GTATCCGGTACTCTGTAGCCGATATAGGTGGCTCACTTCTATATCCGACCGATAGGA 297
QY
101 IleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPheTTPProAlaGly 120
Db
298 ATAATAGGTCTATAATAATATCTTTTGGTACCCTCAATCACTGTCTTTTGGCCCGGGGA 357
QY
121 GluGlnAspLysThrValTTPThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db
358 GAACAAAGACAAAACAGTATGACACAAATTTATTAATGGAGAAATTTTGTGTATACA 417
QY
141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGlyPheArgGlnIle 160
Db
418 CCGTTAACAGAAAGCATAAAACAGTAAAGTTACAAACTTTAGAAAGGATTTAGACAAATA 477
QY
161 LeuGlnSerTyAsnThrAlaLeuAspAspTTPArgLysLeuLysArgLeuGlnAlaPro 180
Db
478 TTACAAAGCTATAATAACAGCATTAGATGATGGAGAAATTTAAAGAGACTACAAAGCTCCT 537
QY
181 GlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGlu 200
Db
538 GGATTAACCAACCATCATCAGATTACAAACAGTGCCTTGACTCTTAAATATACGATTTGAG 597
QY
201 AsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyLysThr 220
Db
598 AATGTTCACAATGATTTTATCGAGAAATACCTGGTTTCCAACTTGAAACTTATATAAACG 657
QY
221 LeuLeuLeuProIleTyAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGly 240
Db
658 CTATTACTACCTATTATGCGCAAGCTGCTTAATTTTCAATTAATTTATTAACAACAGGT 717
QY
241 AlaGluLeuAlaAspGluTTPAsnAlaAspIleHisProSerGlnIleGluProAsnAla 260
Db
718 GCTGAATGCGTGATGAAATGGAATGAGATACATCCCTTCAAAATTTGAACCTAATGCT 777
QY
261 GlyThrSerAspAspTyTyTyLysLeuLysGluAsnIleProLysTySerAsnTy 280
Db
778 GGAACATCAGATGACTATTATTAACCTTTTAAAGAGAAATATACCTAAATATAGTAACAT 837
QY
281 CysAlaAsnThrTyArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTTP 300
Db
838 TGTGCAAAATACCTATAGAGAAGGACTTAAATAAACTTCGAAACGAACTAATATGAGATGG 897
QY
301 SerIlePheAsnAspTyArgArgTyMetThrIleThrValLeuAspThrIleSerGln 320
Db
898 AGTATATTTAATGATTATCGAAGATATAGACTATTACTGTATTAGATACATATCGCTCAA 957
QY
321 PheSerLeuTyAspIleLysArgTyArgAspSerIleGlyGlyIleGluValLysGly 340
Db
958 TTTTCTTTTATGATATAAGAGATCAAGATTCATAGGA-----AGAAATAGGTGGC 1011
QY
341 IleLysAsnGluLeuThrArgGluIleTyThrThrGluIleAsnPheAspArgLeuPro 360
Db
1012 ATTAAACTGAACTTACAAGAGAAATTTATACAACCTGAAATAAATTTTACCCTCTTACT 1071
QY
361 GlnLeuArgValGlnProAsnLeuAlaThrMetGluTyAsnLeuThrArgAlaSerPhe 380
Db
1072 TACCTTGAATTAACCCCAATCTCGCTATAATGGAATATAATTTAAACAGCTTCAGGGCTT 1131
QY
381 LysLeuPheSerPheLeuGluGlnPheIlePheTyThrGlnAsnThrAsnPheGlyAsn 400
Db
1132 AGATTATTTTCACTTTTAGTAGAATCTTATTTATACAAATAATGAAACGTACGGGAAT 1191
QY
401 ArgLeuValGlyIleSerAsnArgAspAlaProThrTySerAsnThrIleThrGluThr 420

Db
1192 CGTTAGTTGGTATTGCAATCGTAATAGATCTACTATTATGCTACGACAGGAAGTGAATTT 1251
QY
421 LeuTyTyGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTy 440
Db
1252 ATATATGGAGAAAGAACACAGGTCCACCACACAAACAACTTTAATACCATTTTGAATCTTAT 1311
QY
441 LysValSerIleValThrAspArgGlnSerProProValSerProIle--GlnProHis 459
Db
1312 AAAAGTTCAATTTGAATGATAGACAAGTAACCTCTACTTCCCTCTTCTTCTTAACATATAC 1371
QY
460 PheIleIleAsnGlnIleGluLeuTyLeuAsnGlySerSerAsnAsnThrLeuLysTy 479
Db
1372 TTTTACAATTAATCAATTTGAATTTTAAATTAATTTACCTAGTAATAAATTAACATAT 1431
QY
480 SerAlaGlyGlySerLeuSerAsnTyGlnAsnThrThrPhePheGlnPheProArgLys 499
Db
1432 TCAGCTGGGGGAATTTATCTAATGATAAAAAACAACTGATTTTCAATTTCTGTAAAA 1491
QY
500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTySerHis 519
Db
1492 AAAGACTGTAAACCAATTTAATCCAAATTTGTTTACCAAGCTATAATAGTTATAGTCAT 1551
QY
520 IleLeuSerHisPheSerLeuPheThrTySerTyValIleGlyLeuGlnGlnIle 539
Db
1552 ATTTTATCCAGTTTCTTTTATTAATTTCTTATAAATTTGGATTAGCCCTAAATATA 1611
QY
540 LeuAspThrGlyValLeuGlyTTPThrHisSerSerValAspArgTyAsnAlaIleSer 559
Db
1612 TTATATACAGTGCATTTAGGATGGACACACAGTAGTGTAAATAGAAATAATGCAATATCA 1671
QY
560 AspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLys 579
Db
1672 GATAAAATAATTAACAATGATCCAGCAATCAAGGTAACAGCTCTGTATACAACTCTAAG 1731
QY
580 ValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyLeuGlnSerGlnGlyArg 599
Db
1732 GTAATTTGAAGAGCTGGTCTACAGGAGAACTTGGTTTATTTACAAAGTCAAGGGCGT 1791
QY
600 LeuGluIleThrCysGluThrProAsnSerThrGlnSerTyPheIleArgLeuArgTy 619
Db
1792 TTAGAGATTACATGTAGAACTCTTAATTTACACAATCTTATTACATTAGACTTCGATAC 1851
QY
620 AlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIle 639
Db
1852 GCTACAAATGGTGTGGAATACTCTTCTTAATATATCTCTTACAATACCAAGGAGTAATA 1911
QY
640 GlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyAsnAsnLeuGln 659
Db
1912 GGAATACCCTCAACGACTCAACAAACACTTTTCTGGTACAAATTAATAATTTTACAA 1971
QY
660 TyTyGlyAspPheGlyTyPheGlnPheProSerThrValThrLeuProLeuAsnArgAsn 679
Db
1972 TAGGAGATTTTGGTATTTCCTTAATTTCCAAAGTACAGTAACATTACCTTTAAATCGAATC 2031
QY
680 IleProPheIlePheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIle 699
Db
2032 ATACATTTTATATTTAATCGTCAGATGTATCAAAATTTCAATTTTAAATCATGATAAAT 2091
QY
700 GluPheIleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThr 719
Db
2092 GAATTTTATACCAATTTACTCTCTGTACGCCAAATAATAGAGAAAAACAAATTTAGAACT 2151
QY
720 IleGlnThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db
2152 ATCAAAACAAAAATAAATACATTTTTCACAAATCATCAAAA 2193

RESULT 3

ADR89398 standard; cDNA; 2085 BP.

XX ADR89398;

AC
XX
DT 18-NOV-2004 (first entry)

XX AXMI-007 alternative start site coding sequence.

DE ss; gene; delta-endotoxin; delta-endotoxin associate polypeptide;
KW expression cassette; transformation; transgenic; plant; bacteria;
KW lepidoptera; coleoptera; pest; pesticide; resistance;
KW pesticidal activity.

XX Bacillus thuringiensis.

XX Key Location/Qualifiers

FT CDS 1. 2085

FT /*tag= a

FT /product= "Alternative AXMI-007"

PN WO2004074462-A2.

XX

PD 02-SEP-2004.

XX 20-FEB-2004; 2004WO-US005829.

XX

PR 20-FEB-2003; 2003US-0448632P.

PR 20-FEB-2003; 2003US-0448633P.

PR 20-FEB-2003; 2003US-0448797P.

PR 20-FEB-2003; 2003US-0448806P.

PR 20-FEB-2003; 2003US-0448810P.

PR 20-FEB-2003; 2003US-0448812P.

PR 19-FEB-2004; 2004US-00781979.

PR 19-FEB-2004; 2004US-00782020.

PR 19-FEB-2004; 2004US-00782096.

PR 19-FEB-2004; 2004US-00782141.

PR 19-FEB-2004; 2004US-00782570.

PR 19-FEB-2004; 2004US-00783417.

XX (ATHE-) ATHENIX CORP.

PA Carozzi N, Hargiss T, Koziel MG, Duck NB, Carr B;

XX WPI; 2004-635574/61.

DR P-PSDB; ADR89399.

XX

PT New isolated delta-endotoxin and delta-endotoxin-associated nucleic acids
PT and polypeptides, useful for killing lepidopteran or coleopteran pests or
PT for producing organisms with pesticide resistance.

XX Claim 1; SEQ ID NO 10; 178pp; English.

XX

CC This sequence encodes an isolated delta-endotoxin. Some of the delta-
CC endotoxin coding sequences of the invention have alternative start
CC codons, producing more than one protein from a single open reading frame.
CC The nucleic acid sequences of the invention are useful in DNA constructs
CC or expression cassettes for transformation and expression in plants and
CC bacteria. The nucleic acids and corresponding polypeptides are useful for
CC killing lepidopteran or coleopteran pests. Compositions containing the
CC delta-endotoxins of the invention, and methods for their production, are
CC useful for the production of organisms with pesticide resistance,
CC specifically bacteria and plants. These organisms are useful for
CC generating altered or improved delta-endotoxin or delta-endotoxin-
CC associated proteins that have pesticidal activity, or for detecting the
CC presence of delta-endotoxin or delta-endotoxin-associated proteins or
CC nucleic acids in products or organisms.

XX Sequence 2085 BP; 794 A; 338 C; 302 G; 651 T; 0 U; 0 Other;

SQ

Alignment Scores:

Pred. No.:	7.9e-271	Length:	2085
Score:	3104.50	Matches:	598
Percent Similarity:	91.95%	Conservative:	30
Best Local Similarity:	87.55%	Mismatches:	52
Query Match:	80.24%	Indels:	3
DB:	13	Gaps:	2

US-10-783-417-2 (1-735) x ADR89399 (1-2085)

Qy	52	MetCysGlnGlyAsnThrGlnTyrGlyAsp	PhenPheGluThrPheAlaSerAla	aspThr 71
Db	1	ATGTGTCAAGGAATACACAATATGGTGA	TATATTCGAGACATTTGCTAGTCTG	TATACA 60
Qy	72	IleAlaValSerAlaGlyThrIleValSer	GlyThrLeuLeuAlaGlyIleGlyGly 91	
Db	61	ATTGCTGCAGTTAGTGCAGGTACTATTG	TATCCGGTACTCTGTTAGCCGGTATAG	TGGTGG 120
Qy	92	LeuThrSerIleSerGlyProIleGlyIle	IleGlyAlaIleIleIleSerPheGly	Thr 111
Db	121	CTCAGTCTATATCCGGACCGATAGGAAT	ATAGTGTCTATAATAATATCTTTTGT	TACC 180
Qy	112	LeuIleThrValPheTyrProAlaGlyGlu	GlnAspLysThrValTyrThrGlnPhe	Ile 131
Db	181	CTAATCACTGTCTTTTGGCCCGCGGAGA	CAAGACAAACAGATATGACACAATTT	TATT 240
Qy	132	LysMetGlyGluIlePheValAspThrPro	LeuThrGluSerIleLysGlnLeuLys	Leu 151
Db	241	AAATGGGAGNAATTTTGTGTATACCCG	TATACAGAAAGCATAAACAGCTAAAG	TTA 300
Qy	152	GlnThrLeuGluGlyPheArgGlnIleLeu	GlnIleSerTyrAsnThrAlaLeuAsp	Thr 171
Db	301	CAAACTTTTAGAAGGATTAGACAAATAT	TACAAAGCTATAATACAGCATTAGAT	TGG 360
Qy	172	ArgLysLeuLysArgLeuGlnAlaProGly	LeuProSerSerAlaLeuGlnAla 191	
Db	361	AGAAATTTAAAGAGACTACAAAGCTCT	CGGATTTACCACCATCATCAGCAAGCT	420
Qy	192	AlaLeuThrLeuLysIleArgPheGluAsn	ValHisAsnAspPheIleArgGluIle	Pro 211
Db	421	GCCTTGACTCTTAAATACGATTTGAGAA	TGTTGACAATGATTTTATTCGAGAAAT	ACCT 480
Qy	212	GlyPheGlnLeuGluThrTyrTyrLysThr	LeuLeuProIleTyrAlaGlnAlaAsn	231
Db	481	GGTTTCCAACTTGAACTTATAAAACGCT	ATTACTACTTTATGCGCAAGCTGCTAT	540
Qy	232	PheHisLeuAsnLeuLeuGlnGlnGlyAla	GluLeuAlaAspGluTyrAsnAlaAsp	Ile 251
Db	541	TTTCATTTAAATTTATTACAAAGGTGCT	GAAATGGCTGATGAATGGAATCAGATA	TATA 600
Qy	252	HisProSerGlnIleGluProAsnAlaGly	ThrSerAspAspTyrTyrLysLeuLys	271
Db	601	CATCCTTTCAACAAATTTGAACCTTAAT	GCTGGAACATCAGATGACTATTATAAA	CTTTAAAA 660
Qy	272	GluAsnIleProLysTyrSerAsnTyrCys	AlaAsnThrTyrArgThrGlyLeuLys	Asn 291
Db	661	GAAATATACCTAATATAGTAATTTGTC	CAATATCTATAGAGAGAGCTAATAATA	720
Qy	292	LeuArgAspGluProAsnMetLysTrpSer	IlePheAsnAspTyrArgTyrMetThr	311
Db	721	CTTCGAAACGAACCTAATATGAGATGGA	TATATTTAATGATTCGAAGATATATG	ACT 780
Qy	312	IleThrValLeuAspThrIleSerGlnPhe	SerLeuTyrAspIleLysArgTyrArg	Asp 331
Db	781	ATTACTGTATTAGATACCTATCGCTCAAT	TTCTTTTATGATATATAAGAGATACA	AAAGAT 840
Qy	332	SerIleGlyGlyIleGluValLysGlyIle	LysAsnGluLeuThrArgGluIleTyr	Thr 351
Db	841	TCATAGGA-----AGATAGGTGGCANT	AAACTGAAGCTTACAGAGAAATTTATA	CATA 894
Qy	352	ThrGluIleAsnPheAspArgLeuProGln	LeuProGlnLeuArgValGlnProAsn	LeuAlaThrMet 371
Db	895	ACTGAATATAATTTTGGACCGTCTTACT	TTACCTTGAAATTCACCAATCTCGCT	ATAATG 954
Qy	372	GluTyrAsnLeuThrArgAlaSerPheLys	LeuPheSerPheLeuGluGlnPheIle	Phe 391
Db	955	GAATATAATTTAACACGTTACGGGCTTA	GTAGATTATTTTCAATTTTATAGATCA	ACTTATATT 1014
Qy	392	TyrThrGluAsnThrAsnPheGlyAsnArg	LeuValGlyIleSerAsnArgAspAla	Pro 411
Db	1015	TATACAAAAATGAACGTCACGGAAATCG	TTTGTAGTTGGTATTGCGAATCGTAAT	AGATCT 1074

Db 1065 GGAGATTTTGAACCTTTTATTGATAGT-----GGTGAACCTCAGTGCCTATACTATT 1115
Qy 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
Db 1116 GTAGTTGGACCGTACTGAGCTGGTTTCGGGTTCACAACA-----CCCTTAGGA 1163
Qy 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGly 120
Db 1164 CTT-----GCTTTAATAGGTTTGGTACATTAATACAGTTCCTTTCCAGCCCAA 1214
Qy 121 GluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db 1215 GACCAATCAACACA-----TGGAGTGACTTTTATACACAAACTAAAAATATTATAAAAAA 1271
Qy 141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
Db 1272 GAAATGACATCAACATATATAGTAATAGTCTAATAAAAAATTTTAAACAGGTCGTTTAATGTT 1331
Qy 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaPro 180
Db 1332 ATCAGCACTTATCATATACCTTAAACATGGAG-----AATATATCCA 1376
Qy 181 GlyLeuProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGlu 200
Db 1377 AACCCACAAATACTCAGGATGTAAGGACACAAATCCAGCTAGTTCATTACCAATTTTCAA 1436
Qy 201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr 218
Db 1437 AATGTCAATCCAGAGCTTGTAAACTCTTGTCTCTAATCTAGTCTAGTATGCGATTACTAT 1496
Qy 219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGln 238
Db 1497 AACATACAGTATTATCTAGTATTGCAACAGCAGCAAACTTACATCTGACTGTATTAAT 1556
Qy 239 GlnGlyAlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluPro 258
Db 1557 CAAGCCGTCAAAATTTGAAGCGTATTTAAACAAATCGCAATTCGATTATTAGAGCCT 1616
Qy 259 AsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSer 278
Db 1617 ---TTGCCAACAGCAATTTGATTATATATCCAGTATTGACTAATGCTATAGAGATTACACT 1673
Qy 279 AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro----- 296
Db 1674 AATTATTGTGTAAACAACCTTATAAAAAAGGATTAATTTAATTAACACGACCCCTGATAGT 1733
Qy 297 -----AsnMetLysTyrSerIlePheAsnAspTyrArgTyrMetThrIle 312
Db 1734 AATCTTGATGGAATATAAAGTGAACACATACATACATAGTATCGAATCAACAAATGACTACT 1793
Qy 313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer 332
Db 1794 GCTGTATTAGATGTTGTGCACTCTTCTCTAATTAATGATGATGATGATTAAT- 1844
Qy 333 IleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThr 352
Db 1845 -----CCAATAGGTGTCCAATCTCAACTTACTCTGAGAAAATTTATCAGGTA 1899
Qy 353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
Db 1890 ---CTTAACCTCGAAGAAAGCCCTATAAATATTAT-----GACTTTCATATCAAGAG 1940
Qy 373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyr 392
Db 1941 GATTCACTTACAGTAGACCG---CAATTATTACTTGGCTTGATCTTGAATTTTAT 1997
Qy 393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
Db 1998 GAAAAAGCGCAAACTACTCTCTAATAATTTTTCACGACCACTATATAATATGTTTCATTAC 2057
Qy 402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
Db 2058 ACACCTTGATAATATATCCAAAAATCTAGTGTGTTTTTGGAAATCACAATGTAACCTGATAAA 2117

RESULT 5

AAQ14669

ID AAQ14669 standard; DNA; 3543 BP.

XX

Qy 421 LeuTyrGlyGluArgThrGlySerProThrThrThrLysThrIleArgProPheGluSerTyr 440
Db 2118 TTA-----AAATCTCTTGGTTTGGCAACAATATTATATATT----- 2153
Qy 441 LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
Db 2154 -----TTT 2156
Qy 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
Db 2157 TTATTAAATGTATAGCTTAGATAATAAATATCTAAATGATTATAATAATATTAGTAAA 2216
Qy 477 -----LeuLysTyrSerAlaGly 483
Db 2217 ATGGATTTTTTATAACTAATGTTACTAGACTTTTGGAGAAAGAACTTACAGCAGGATCT 2276
Qy 484 SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysLysAspCys 502
Db 2277 GGGCAAAATACTTATGATGTAATAATAAATATTTTCGGGTTACCAATCTTAAACGAAGA 2336
Qy 503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
Db 2337 GAGAAATCAAGGAACCCCTACCCCTTTTCCACATATGATACTATAGTTCATATTTATCA 2396
Qy 523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuIleLeuAspThr 542
Db 2397 TTTATTAAAGTCTTAGTATCCCTGCAACATATAAACTCAAGTGTAT----- 2444
Qy 543 GlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
Db 2445 ---ACGTTTGTGTGACACACTCTAGTGTGATCTATAAAATACAAATTTATACACATTTA 2501
Qy 563 IleThrMetIleProAlaIleLysGlyAsnLeuAspThrAsnSerLysValIleGlu 582
Db 2502 ACTACCCAAATTCAGCTGTAAAGCGAATTCACCTGGGACTGCTTCAAGGTTGTTCAA 2561
Qy 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle 602
Db 2562 GGAGCTGTGTCATACAGGAGGGATTTAAT-----GATTTCAAAAGATCATTTCAAAAT 2615
Qy 603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
Db 2616 ACATGTCAACACTCAAAATCTTCAACAATCGTATTTTATAAGAAATTCGTTATGCTTCAAT 2675
Qy 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db 2676 GGAAGCGCAATATCTCGAGCTGTTATAATCTTAGTATCCAGGGGTAGCAGAACTG--- 2732
Qy 643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
Db 2733 GGTATGGCACTCAACCCCACTTTTCTGTCGACAGATTATACGAATTTAAATATAAAGAT 2792
Qy 663 PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
Db 2793 TTTCACTACTTAGAATTTTCTAAACGAGGTGAAATTTGCTCCAAATCAAAACATATCTCT 2852
Qy 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleAspLysIleGluPhe 701
Db 2853 GTGTTTAATCTGTCGATGATATATAACAAACACACAGACTTATTATTGATAAAATTTGAAT 2912
Qy 702 IleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGln 721
Db 2913 CTGCCAATTACTCGTCTTATAAGAGAGATAGAGAGAAACAAATAATTAGAACAGTACAA 2972
Qy 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db 2973 CAATAATTAATACATTTTATGCAAAATCCTATAAAA 3008


```
Db 274 CTT-----GCTTTAATAGCTTTTGGTACATTAATACCAGTCTCTTTTTCAGCCCAA 324
Qy 121 GluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyLeuIlePheValAspThr 140
Db 325 GACCAATCTAACACA---TGGAGTGTCTTTATACACAACTAAATATATATAAAAAA 381
Qy 141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
Db 382 GAAATAGCATCAACATATATAAGTAATGCTAATAAATTTTAAACAGGTGCTTTAATGTT 441
Qy 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaPro 180
Db 442 ATCAGCACTTATCATAAATCACCTTAAACATGGGAG-----AATAATCCA 486
Qy 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGlu 200
Db 487 AACCCACAAAATACTCAGAGTGTAGGACACAAATCCAGCTAGTTTCATTACCAATTTCAA 546
Qy 201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr 218
Db 547 ATGTGATTCAGAGCTTGTAACTCTTGTCTCTTAATCCTAGTGTATTCGATTAAT 606
Qy 219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGln 238
Db 607 AACATACTAGTATTATCTAGTTATGCAAGCAGCAAACTTACATCTGACTGTATTAAAT 666
Qy 239 GlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluPro 258
Db 667 CAAGCCGTCAAAATTTGAACGCTATTTAAAAACAATCGACAATTCGATTATTAGAGCCT 726
Qy 259 AsnAlaGlyThrSerAspAspTyrTyrLysLeuLysGluAsnIleProLysTyrSer 278
Db 727 ---TTGCCAACGAATTCGATTATATCCAGTATTCAGTAAGCTATAGAGATTACACT 783
Qy 279 AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro----- 296
Db 784 AATTATTGTGAACAACCTTATAAAAAAGGATTAAATTTAATTAACAGCGCTGTAGT 843
Qy 297 -----AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIle 312
Db 844 AATCTTGATGGAATATAAATCGGAACACATACAATACGTATCGAACAATAATGACTACT 903
Qy 313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer 332
Db 904 GCTGATTAGATCTGTTCGACTCTTCTCTAATATATGATGATGTAATAT----- 954
Qy 333 IleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThr 352
Db 955 -----CCAATAGGTGTCCTAATCTGAACCTTACTCGAGAAATTTATCAGGTA 999
Qy 353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
Db 1000 ---CTTAACCTCGAAGAAAGCCCTATAAATATTAT-----GACTTTCAATATCAAG 1050
Qy 373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyr 392
Db 1051 GATTCACTTACAGTACGACG---CAATTATTACTTGGCTTGATCTTTGGAATTTTTAT 1107
Qy 393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
Db 1108 GAAAAGCGCAAACTACTCTCTAATAATTTTTCACAGCCATTAATAATGTTTCATTAC 1167
Qy 402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
Db 1168 ACACCTTGATAATATATCCCAAAAATCTAGTGTCTTTTGGAAATCACAATGTAACGTATAA 1227
Qy 421 LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440
Db 1228 TTA-----AAATCTCTTGGTTTGGCAACAAATATTATATATT----- 1263
Qy 441 LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
|||
```

```
Db 1264 ----- 1266
Qy 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
Db 1267 TTATTAATCTCATAGCTTAGATAATAAATATCTAATATGATTATAATAATATTAGTAAA 1326
Qy 477 -----LeuLysTyrSerAlaGlyGly 483
Db 1327 ATCGATTTTTTTTATAACTAATAGTGTACTAGACTTTTGGAGAAAGAACTTACAGCAGGATCT 1386
Qy 484 SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysLysAspCys 502
Db 1387 GGGCAATAACTATTATGATGATAAATAAAYATTTTCGGGTACCAATTTCTTAACCAAGA 1446
Qy 503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
Db 1447 GAGATCAAGCAATCCCTACCTTTTCCACATATGATATACTATATGTCATATTTTATCA 1506
Qy 523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThr 542
Db 1507 TTTATTAAAAAGTCTTAGTATCCCTGCAACATATAAAACTCAAGTGTAT----- 1554
Qy 543 GlyValLeuGlyTrpThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
Db 1555 ---ACGTTTCTGCTGGACACACTCTAGTGTGTGATCTCTAAAAATACAAATTTATACACATT 1611
Qy 563 IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582
Db 1612 ACTACCAAAATTCAGCTGTAAAGCGAATTCACCTGGGACTGCTTCTTAGGTGTTCNA 1671
Qy 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle 602
Db 1672 GGACCTGTCATACAGGAGGGGATTTAAT-----GATTTCAAAGATCATTTCAAAAT 1725
Qy 603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
Db 1726 ACATGTCAACACTCAAAATTTTCAACAATCGTATTTTATAAGAAATTCGTATGCTTCAAA 1785
Qy 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db 1786 GGAAGCCCAATACAGAGCTGTATTAATCTTAGTATCCAGGGGTAGCAGAACTG--- 1842
Qy 643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
Db 1843 GGTATGCACTCAACCCCACTTTTCTGTCAGATATATACGAATTTAAATATAAAGAT 1902
Qy 663 PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
Db 1903 TTTCACTACTTGAATTTTCTTAACGAGGTGAAATTTGCTCCAAATCAAAACATATCTCTT 1962
Qy 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleAspLysIleGluPhe 701
Db 1963 GTGTTTAATCGTTCGGATGATATATACAAACACACAGTACTTATTGATAAAATTTGAATTT 2022
Qy 702 IleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGln 721
Db 2023 CTGCCAATTAATCTCTTCTATAGAGAGGATAGAGAGAAACAAATATAGAACAGTACAA 2082
Qy 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db 2083 CAATAATAATTAATCATTTTATGCAAAATCTCTATAAAA 2118
RESULT 7
AAN93054
ID AAN93054 standard; DNA; 3940 BP.
XX
AC AAN93054;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 30-MAY-1990 (first entry)
XX
DE Delta-endotoxin crystal protein gene.
```

1164	Db	CTT-----GCTTTAATAGGTTTGGTGACATTAATACCACTTCTTTTCCAGCCCAA	1214
121	Qy	GlulGlnAspLysThrValTThrGlnPheIleLysMetGlyGlulPheValAspThr	140
1215	Db	GACCACTCTAACACA---TGGAGTGACTTTATAACACAACAATAATATATAAAAAA	1271
141	Qy	ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle	160
1272	Db	GAATAGCATCAACATATATAAGTAATGCTAATAAAAAATTTTAAACAGGTCGTTAATGTT	1331
161	Qy	LeuGlnSerTyrAsnThrAlaLeuAspAspTTPArgLysLeuLysArgLeuGlnAlaPro	180
1332	Db	ATCAGCACTTATCATATATCACTTAAACATGGAG-----AATAATCCA	1376
181	Qy	GlyLeuProProSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGlu	200
1377	Db	AACCCACAAATATCTCAGGATGTAAGACACAATCCAGCTAGTCTCATACCAATTTTCAA	1436
201	Qy	AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr	218
1437	Db	AATGTCATCCACAGCTTGTAAACTCTTGTCCTCTTAATCTAGTATTGCGATTACTAT	1496
219	Qy	LysThrIleLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGln	238
1497	Db	AACATATAGTATTATCTAGTTATGCAACAGCAGCAAACTTACATCTGACTGTATTAAAT	1556
239	Qy	GlnGlyAlaGluLeuAlaAspGluTTPAsnAlaAspIleHisProSerGlnIleGluPro	258
1557	Db	CAGCCGCTCAAAATTTGAAGCGTATTTAAAAACAATCGCAATTCGATTTATTAGAGCCT	1616
259	Qy	AsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSer	278
1617	Db	---TTGCCAACGCAATTGATTTATTCAGTATTGCTTAAAGCTATAGAGATTACACT	1673
279	Qy	AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro-----	296
1674	Db	AATTATTGTGTACAACTTATAAAAAAGGATTTAAATTTAAATAAACGCGCTGATAGT	1733
297	Qy	-----AsnMetLysTPSerIlePheAsnAspTyrArgArgTyrMetThrIle	312
1734	Db	AATCTTGATGGAAATATAACTCGAACACATACATGATCATCGAACAAAATGACTACT	1793
313	Qy	ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer	332
1794	Db	GCCTGATTAGATGTTCTGCACCTCTTCTCTTAATTATGATGTAGTAAATAT-----	1844
333	Qy	IleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThr	352
1845	Db	-----CCAAATAGGTGTCCAATCTCGAACTTACTCGAGAAATTTATCAGGTA	1889
353	Qy	GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu	372
1890	Db	---CTTAACCTCGAAGAAAGCCCTATAAATATTAT-----GACTTCAATATCAAGAG	1940
373	Qy	TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyr	392
1941	Db	GATTCATCTACAGTAGACCG---CATTTATTACTTGGCTTGATCTTTGAATTTTAT	1997
393	Qy	-----ThrGluAsnThrAsnPheGlyAsnArg-----	401
1998	Db	GAATAAAGCGCAACTACTCTATAATAATTTTTCACCAGCCATTTATAATATGTTTCATTAC	2057
402	Qy	---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr	420
2058	Db	ACACTTGATAATATATCCCAAAATCTAGTGTTTTGGAAATCACAATGTAACACTGATAA	2117
421	Qy	LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr	440
2118	Db	TTA-----AAATCTCTTGGTTGGCAACAATATTTATTT-----	2153
441	Qy	LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe	460

PN JP2004166574-A.
 XX 17-JUN-2004.
 XX 19-NOV-2002; 2002JP-00335440.
 XX 19-NOV-2002; 2002JP-00335440.
 XX (UYOK-) UNIV OKAYAMA.
 XX WPI; 2004-445549/42.
 XX Novel codon-modified gene encoding Cry4A protein, the modification being
 PT substitution with most frequently used codons in host microorganisms such
 PT that resulting gene has high translation rate in the microorganism.
 XX Claim 3; SEQ ID NO 1; 20pp; Japanese.
 XX The invention comprises a codon-modified gene which codes for Cry4A
 CC protein, the gene is modified by substitution with the most frequently
 CC used codon in the host microorganism, so that the modified gene will have
 CC a high translation rate in the host microorganism. The codon-modified
 CC gene of the invention is useful for the production of Cry4A protein in
 CC large quantities. The present DNA sequence represents a codon-modified
 CC Cry4A gene of the invention.
 XX Sequence 2130 BP; 677 A; 704 C; 428 G; 321 T; 0 U; 0 Other;
 SQ

Alignment Scores:
 Pred. No.: 3,64e-83 Length: 2130
 Score: 1032.50 Matches: 265
 Percent Similarity: 50.26% Conservative: 117
 Best Local Similarity: 34.87% Mismatches: 273
 Query Match: 26.69% Indels: 105
 DB: 12 Gaps: 24

US-10-783-417-2 (1-735) x ADP71294 (1-2130)

QY 1 MetAsnGlnAsnAspAsnAsnGluTyrGluIleAspSerHisThrSerProTyr 20
 DB 30 ATGAACCCGTACCAAAACGAAGACGAAATACGAACCCCTGAACCGCGCAGAAAGAACTG 89
 QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
 DB 90 -----AACATCAGCAACAACTACACCCGTTACCGATCGAAGAACGCCGGAACAACTG 143
 QY 41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
 DB 144 CTGCAAGCACCACCACTACAAAGACTGGCTGAACATGTGCCAACAAACCAACCAATACGGC 203
 QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80
 DB 204 GCGCACTTCGAACCTTCATCGACAGC-----GGTGAACGTGACGCGGTACACCATC 254
 QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
 DB 255 GTGGTCGACCGCGTGTACCGGTTTCGGCTTCACACAC-----CCGCTGGGC 302
 QY 101 IleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGly 120
 DB 303 CTG-----CGCTGATCGGTTTCGGTACCTCGGTGCTGTTCCCGGCGCAA 353
 QY 121 GluGlnAspLysThrValTyrThrGlnPheIleIleMetGlyGluIlePheValAspThr 140
 DB 354 GACCAAGCAACACAC-----TGGAGCGACTTCATCACCCCAACCAACCAACATCATCAAAAA 410
 QY 141 ProLeuThrGluSerIleLysGlnLeuLysGlnThrLeuGluGlyPheArgGlnIle 160
 DB 411 GAAATCGCAAGCAGCTACATCAGCAACGCGCAACAAATCTGAAACCGTAGCTCAACGTG 470
 QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaPro 180
 DB 471 ATCAGCACCTACCAACACCACTGAAACCTGGAA-----AACCAACCCG 515

QY 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGlu 200
 DB 516 AACCCGCAAAACACCCAAAGAGTGGTACCCAAATCCAACTGGTGCGTACCACTTCCAA 575
 QY 201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr 218
 DB 576 AACGTCAATCCCGGAACCTGGTGAACAGCTGCCCGCGCAACCCGAGCGAGCTGCGACTAC 635
 QY 219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGln 238
 DB 636 AACATCTGGTGTGACGAGCTACCGCAAGCAGCAGCACTGCACTGACCTGACCGTGTGAAC 695
 QY 239 GlnGlyAlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluPro 258
 DB 696 CAAGCGGTCAAAATTCGAGGCTTACCTGAAACCAACCGCAATTCGACTACCTCGAGCGG 755
 QY 259 AsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSer 278
 DB 756 ---CTGCCGACCGCAATTCGACTACTACCGGTGCTGACCAAGCAATCGAAGACTACACC 812
 QY 279 AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro----- 296
 DB 813 AACTACTGCTGACCACTACAAAAAGGCTGAACCTGTATCAAAACCAACCCCGGACAGC 872
 QY 297 -----AsnMetLysTrpSerIlePheAsnAspTyrArgTyrMetThrIle 312
 DB 873 AACTGACGCGCAACATCACTGGAACACCTACACCACTACCGCAACCAATGACCAACC 932
 QY 313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer 332
 DB 933 CGGTGTGCTGACCTGTGGCACTGTTCGCAACTACGACGTGCGGTAAATAC----- 983
 QY 333 IleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThr 352
 DB 984 -----CCGATCGGTGTCCAAAGCAACTGACCCGGGAATCTACCAAGTC 1028
 QY 353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
 DB 1029 ---CTGAACCTCGAAGAACCGCTACAAATACTAC-----GACTTCCAATACCAAGAA 1079
 QY 373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyr 392
 DB 1080 GACAGCTGACCGCTGCGCCG---CACGTGTTCACCTGGCTGGACAGCTGAATCTCTAC 1136
 QY 393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
 DB 1137 GAAAAAGCGCAACACCCCGCAACAACTTCTTACCGCCCACTACAAATGTTCCACTAC 1196
 QY 402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
 DB 1197 ACCCTGACCAACATCAGCAAAAGAGCGGTGTCGCAACCAACACGTCGACGCAAA 1256
 QY 421 LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440
 DB 1257 CTG-----AAAAGCCTGGGTCTGCAACCAACATCTATCATC----- 1292
 QY 441 LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
 DB 1293 -----TTC 1295
 QY 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
 DB 1296 CTGCTGAACGTCACTAGCTGGACAAACAATACCTGAACGAGCTACAAACATCAGCAAA 1355
 QY 477 -----LeuLysTyrSerAlaGlyGly 483
 DB 1356 ATGGAATCTTTCATCACCAACGCTACCGCTGCTGCAAAAGAACTGACCGGCGCAGC 1415
 QY 484 SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysLysAspCys 502
 DB 1416 GGCCAAATACCTACGACGTGAACAAACAAATCTTTCGGCTGCTCCGATCTCTGAACGTCGC 1475

```
QY 503 AsnLeuValIleAseProAsnProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
Db 1476 GAAACCAAGCACCACCGACCTGTTCCGACCTACGACACTACGACCACTCTGAGC 1535
QY 523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThr 542
Db 1536 TTCATCAAAAGCCTGAGCACCCTCCGGCAACCTACAAAACCCCAAGGTAC----- 1583
QY 543 GlyValLeuGlyThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
Db 1584 ---ACCTTCGCGTGGAGCCACAGCAGCGTGCACCCGAAACACCACTCTACACCCACTG 1640
QY 563 IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582
Db 1641 ACCACCAATCCCGCGGTGAAGCAACAGCTGGCCACCGCAGCAAGTGTGCA 1700
QY 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle 602
Db 1701 GGTCCGGGTACACCGCGCGGTGACCTGATC-----GACTTCAAGACCACTTCAAAATC 1754
QY 603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
Db 1755 ACCTGCCAACACACAGCACTTCCAAACAAAGCTACTTCCTCGCATCTCGTACGCGAGCAAC 1814
QY 623 GlyValGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db 1815 GGACGCCAACACCCCGCGGTGATCACTGAGCATCTCCGCGTGGCAGAACTG--- 1871
QY 643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
Db 1872 GGTATGCACTGAACCGGACCTTCAGCGGTACCGACTACCACTGAAATACAAAGAC 1931
QY 663 PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
Db 1932 TTCATATACCTGGAGTTTCAGCAACGAAGTGAATTCGCGCGAACCACAAACATCAGCCTG 1991
QY 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleAspLysIleGluPhe 701
Db 1992 GTTTCACCGTAGCGAGGTATACCAACACCAACCGTGTGTGATCGACAAATCGAGTTC 2051
QY 702 IleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGln 721
Db 2052 CTGCGGATACCCGTAGCATCCGGAAGACCGTGAATAACAAACAACTGGAAACCGTGCA 2111
RESULT 10
ADP71295
ID ADP71295 standard; DNA; 2246 BP.
XX
AC ADP71295;
XX
DT 09-SBP-2004 (first entry)
XX
XX Codon-modified Cry4A gene #2.
XX
XX codon optimisation; codon-modified; Cry4A; gene; ds.
XX
XX Synthetic.
OS Unidentified.
XX
XX JP2004166574-A.
XX
XX 17-JUN-2004.
XX
XX 19-NOV-2002; 2002JP-00335440.
XX
XX 19-NOV-2002; 2002JP-00335440.
XX
XX (UYOK-) UNIV OKAYAMA.
XX
XX WPI; 2004-445549/42.
XX
XX Novel codon-modified gene encoding Cry4A protein, the modification being
PT substitution with most frequently used codons in host microorganisms such
```

```
PT that resulting gene has high translation rate in the microorganism.
XX
PS Disclosure; SEQ ID NO 2; 20pp; Japanese.
XX
CC The invention comprises a codon-modified gene which codes for Cry4A
CC protein, the gene is modified by substitution with the most frequently
CC used codon in the host microorganism, so that the modified gene will have
CC a high translation rate in the host microorganism. The codon-modified
CC gene of the invention is useful for the production of Cry4A protein in
CC large quantities. The present DNA sequence represents a codon-modified
CC Cry4A gene of the invention.
XX
```

Sequence 2246 BP; 711 A; 728 C; 457 G; 350 T; 0 U; 0 Other;

Alignment Scores:

Pred. NO.:	3.92e-83	Length:	2246
Score:	1032.50	Matches:	265
Percent Similarity:	50.26%	Conservative:	117
Best Local Similarity:	34.87%	Mismatches:	273
Query Match:	26.69%	Indels:	105
DB:	12	Gaps:	24

US-10-783-417-2 (1-735) x ADP71295 (1-2246)

```
QY 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 146 ATGAACCGGTACCAAAACAAAGACGAATACGAAACCCCTGGAACGCGCAGCAAGAACTG 205
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 206 -----AACATCAGCACTACACCGCTTACCGATCGAANAACGCCGCAACAACTG 259
QY 41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
Db 260 CTGCAAGACCACTACAAAGACTGGCTGAACATGTGCCAACAAACCAACAATACGCGC 319
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIle 80
Db 320 GCGCACTTCGAACCTTCATCGACG-----GGTGAACCTGAGCGGTACACCATC 370
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
Db 371 GTGTGCGCACCGGTGACCGGTTTCGCTTCCACC-----CGCTGGGC 418
QY 101 IleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGly 120
Db 419 CTG-----GCGCTGATCGTTTCGTATCCCTGATCCCGTCTGTTCGCGGCA 469
QY 121 GluGlnAspLysThrValThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db 470 GACCAAGCAACACCT--TGAGCGACTTCATCACCACCAACCAAAACATCATCAAAAAA 526
QY 141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
Db 527 GAAATCGCAAGCACTTACATCAGCAACGCAACAAATCTCGAACCGTAGCTTCAACGCG 586
QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaPro 180
Db 587 ATCAGCACTTACCAACCACTGTAACCTGGGA-----AACAAACCG 631
QY 181 GlyLeuProProSerSerAlaLeuGlnAlaLeuThrLeuLysIleArgPheGlu 200
Db 632 AACCCGCAAAACACCCCAAGACGTGCTACCCCAATCCAACTGGTGCATCTACCACTTCAA 691
QY 201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr 218
Db 692 AACGTCATCCCGGAACCTGGTGAACAGCTGCCCGCAACCCCGAGGACTGCGACTACTAC 751
QY 219 LysThrLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGln 238
Db 752 AACATCTCTGGTGGTACGAGCTACGCGAAGCAGCAACCTGACCTGACCTGCTGCTGAAC 811
QY 239 GlnGlyAlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluPro 258
```

```
Db      812  CAAAGCGTCAATTCAGGCTTACCTGAAACCAACCGCCAAATTCGACTACCTCGAGCCG 871
Qy      259  AsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSer 278
Db      872  ---CTCCGACCGCAATCGACTACTACCGGTGCTGACCAAGCAATCGAAGACTACACC 928
Qy      279  AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro----- 296
Db      929  AACTACTGGGTGACCACTACAAAGGCGCTGAACCTGATGATAAACCACCCCGGACAGC 988
Qy      297  -----AsnMetLysTyrTrpSerIlePheAsnAspTyrArgArgTyrMetThrIle 312
Db      989  AACCTGGACGGCAACACTGAACACACTCAACACACTTACCGCAACCAAAATGACCAACC 1048
Qy      313  ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer 332
Db      1049  CGGTGCTGACCTGGTGGCAGTGTCCGAACTACGACGTCCGTAAATAC----- 1099
Qy      333  IleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThr 352
Db      1100  -----CGATCGGTGTCCAAAGCGAACTGACCCGCGAAATCTACCAAGTC 1144
Qy      353  GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
Db      1145  ---CTGAACCTCGAAGAAAGCGGTACAAATACTAC-----GACTTCCAAATACCAAGAA 1195
Qy      373  TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyr 392
Db      1196  GACAGCGCTGACCGCTGCGCGG---CACCTGTTTCACTGGCTGGACACGCTGAACCTTCTAC 1252
Qy      393  -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
Db      1253  GAAAGACGCAACACCCCGCAACACTTCTTACAGCCCACTACAACATGTTCCACTAC 1312
Qy      402  ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
Db      1313  ACCCTGGCAACATCAGCCAAAGAGCGGTGTTCGGCAACCAACGACGTCGCGACAAA 1372
Qy      421  LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440
Db      1373  CTG-----AAGAGCTGGGTCTGGCAACCAACATCTACATC----- 1408
Qy      441  LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
Db      1409  -----TTC 1411
Qy      461  IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
Db      1412  CTGCTGAACGTATCATCAGCTGGACAAACAATACTTGAACGACTACAACAACATCAGCAAA 1471
Qy      477  -----LeuLysTyrSerAlaGlyGly 483
Db      1472  ATGGACTTCTTCATCAACACGCTACCGGCTGCTGGAAAGAAAGAACTGACCGGAGCAGC 1531
Qy      484  SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysLysAspCys 502
Db      1532  GGCCAAATCACCTTACGACGTGGAACAAACATCTTCGGCTGCGCATCTGGAACAGTCGC 1591
Qy      503  AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
Db      1592  GAAACCAAGGCACCGGACCTGTTCCCGACCTACGACAACTACAGCACTCTCTGAGC 1651
Qy      523  HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThr 542
Db      1652  TTCATCAAAAGCGCTGAGCATCCCGCAACCTACAAACCCCAAGTGTAC----- 1699
Qy      543  GlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
Db      1700  ---ACCTTCGCGTGGACCCACAGCAGCTGACCCGAAACCAACCACTTACACCCACCTG 1756
Qy      563  IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582
```

```
Db      1757  ACCACCCAAATCCCGCGGTGAAAGCGAAGAGCCTGGGCACCGGAGCAAAAGTGTCCAA 1816
Qy      583  GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle 602
Db      1817  GSTCCGGTTCACACCGCGGTGACCTGATC-----GACTTCAAAGACCACCTTCAAATC 1870
Qy      603  ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
Db      1871  ACCTGCCAACACACAGCAACTTCCAAAGAGTACTTTCATCCGCATCCGTTAGCGAGCAAC 1930
Qy      623  GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db      1931  GGCAGCGCAACACCCGCGGTGATCAACCTGAGCATCCCTGGCGTGGCAGAACTG--- 1987
Qy      643  ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
Db      1988  GGTATGGCACTGAACCGACCTTCAGCGGTACCGACTACCAACCTGAAATACAAAGAC 2047
Qy      663  PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
Db      2048  TTCCAATACCTGGAGTTTCAGCAACGAAGTGAATTCGGCGCGAACCACCAACATCAGCCTG 2107
Qy      683  IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleIleAspLysIleGluPhe 701
Db      2108  GTGTTCAACCGTAGCGAGCTGTACCAACACACACCGCTGCTGATCGACAAATCGAGTTC 2167
Qy      702  IleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGln 721
Db      2168  CTGCGCATCACCGTAGCATCCGCGAAGACCGGTGAAAACCAAAACCTGGAACCGTGCA 2227
RESULT 11
AAQ14670
ID      AAQ14670 standard; DNA; 2061 BP.
AC      AAQ14670;
XX      27-AUG-2003 (revised)
DT      25-MAR-2003 (revised)
DT      04-FEB-1992 (first entry)
XX      Dipteran active toxin gene.
DE      Insecticide; B.t; crystal; delta endotoxin; cryIVC; ss.
KW      Bacillus thuringiensis serovar morrisoni.
XX      Key Location/Qualifiers
PH      CDS 1..2028
FT      /*tag= a
XX      EP457498-A.
PN      21-NOV-1991.
XX      09-MAY-1991; 91EP-00304180.
XX      15-MAY-1990; 90US-00524255.
XX      01-OCT-1990; 90US-00590903.
XX      (MYCO ) MYCOGEN CORP.
XX      Sick AJ;
XX      WPI; 1991-341902/47.
DR      P-PSDB; AARI4374.
XX      Bacillus thuringiensis genes encoding diptera-active toxins - and
PT      transformed microbes used to control insects in various environments.
XX      Claim 1; Page 15; 20pp; English.
XX      The sequence was obtd. from plasmid pMYC1636 which was isolated from a
CC      genomic library prepd. from DNA from B.t. PS71M3 [from B.t. PS71M3-69
```

CC (NRRL B-18515)]. It is related to the cryIVC from B.t. var. israelensis.
 CC The gene encodes a 77 kD protein. Microorganisms transformed with the DNA
 CC may be administered to dipteran insects or their environments, the
 CC expressed toxins acting as an insecticide. See also AAQ14669-Q14672.
 CC (Updated on 25-MAR-2003 to correct PA field.) (Updated on 27-AUG-2003 to
 CC correct OS field.)
 XX

SQ Sequence 2061 BP; 779 A; 331 C; 319 G; 632 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.86e-71	Length:	2061
Score:	899.50	Matches:	247
Percent Similarity:	48.32%	Conservative:	126
Best Local Similarity:	31.99%	Mismatches:	270
Query Match:	23.25%	Indels:	129
DB:	2	Gaps:	30

US-10-783-417-2 (1-735) x AAQ14670 (1-2061)

QY	1	MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr	20
DB	1	ATGAATCCATATCAAAATAGATGAATATGAATATTCATATGCTCCATCCAAATGGTTTT	60
QY	21	PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro	40
DB	61	-----AGCAAGCTTAATACTATTCTAGATATCCATTAGCAATAAGCCAAATCAACCA	114
QY	41	LeuGlnAsnThrAsnTyrIysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly	60
DB	115	CTGAAAACACGAATTACAAAGATTGGCTCAATGTGTCAAGATAATCAACAAATATGGC	174
QY	61	AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIle	80
DB	175	AATAATCGCGGGAAATTTTGTAGTTCTTGAACTATTCTTGAGTTAGTCAGGATATTATT	234
QY	81	ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly	100
DB	235	GTAGTAGAACTATGTAGGA-----GCTTTGTGCTGCCCT-----	270
QY	101	IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGly	120
DB	271	GTCTAGCTGCAGGTATAATATCTTTTGGGACTTTGTTCGGATCTTTTGG---CAAGGA	327
QY	121	GluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr	140
DB	328	TCTGACCCTGCAAATGTTTGGCAGGATTTCGTAACATCGA-----GGAAGG	375
QY	141	ProLeuThrGlu-----SerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPhe	157
DB	376	CCATATCAAGAAATAGATAAAACATAAATTAATGTACTAACTTCTATCGTAACACCTATA	435
QY	158	ArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrPArgLysLeuLysArgLeu	177
DB	436	AAAAATCAACTTGATAAATATCAAGAAATTTTCGATAAATGGGAGCCAGCAGCTF-----	489
QY	178	GlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIle	197
DB	490	-----ACACAGCTTAATCTTAAGACGATACATGATCTCTTTACT	528
QY	198	ArgPheGluAsnValHisAsn-----AspPheIleArgGluIleProGlyPheGln	214
DB	529	ACCTTAGAACCTATAATAGATAAAGATTTAGATATGTTAAATAATAATGCTAGCTATCGA	588
QY	215	LeuGluTyrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeu	234
DB	589	ATACCAACA-----CTCCCTGCATATGCAAAATAGCTACTTGTGCACCTG	633
QY	235	AsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTyrPAsnAlaAspIleHisProSer	254
DB	634	AATTTATTAACATCGCTGCTACCTATTACATATATGG-----CTGCAAAATCAA	684
QY	255	GlnIleGluProAsnAlaGlyThrSerAspTyrTyrLys---LeuLeuLysGluAsn	273

DB	685	GGTATAAATCCAGTACTTTCAATTTCAATTAATTAATCAATCAATCAATCAATCAATCAAT	744
QY	274	IleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArg	293
DB	745	ATCAAGAATATCTACTACTATTGTATACAAACGTACAAATGCAGACTACTATGATTAGA	804
QY	294	AspGluProAsnMetLysTyrSerIlePheAsnAspTyrArgArgTyrMetThrIleThr	313
DB	805	ACTAATACTAACGCAACATGAATATGATATATATCTACCGTTTAGAATGACTCTACT	864
QY	314	ValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIle	333
DB	865	GTGTTAGATCTTATTGCTATTTCCTCAAAATATGACCCAGCAAAATAT-----	912
QY	334	GlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrGlu	353
DB	913	-----CCAATAGGAGTTAAATCTGAACCTTACAGAGAAAGTTTAT---ACGAAT	957
QY	354	IleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyr	373
DB	958	GTTAATTCAGATACATTT-----AGAACCATACAGAACTAGAAAAT	999
QY	374	AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr	393
DB	1000	GGATTAACATAG---AATCTCATATTATTCTTGATAAACCAAGGGCGTTTACACA	1056
QY	394	GluAsnThrAsn-----PheGlyAsnArgLeu	402
DB	1057	AGAAATTCGAGACATTTCTTGATCTTATGATATTTTCTTTTACAGGTAAACAGATG	1116
QY	403	ValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyr	422
DB	1117	-----GCCTTTACATATAATGATGATCGCAACATAATCTCG	1155
QY	423	GlyGluArgThrGly-----SerProThrThrLysThrIleArgProPheGluSer	439
DB	1156	GGACGGTTTCATGGACATATTATTCTCAAGACATCCAAAGTATTTCTTTTATAGA	1215
QY	440	Tyr-----LysValSerIleValThrAspArgGlnSerProProValSerPro	455
DB	1216	AACAAACCTATTGATAAGTCGAAATGTGCAGACATAGAGAGTACTCAGATATAATA---	1272
QY	456	IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn	475
DB	1273	-----TATGAATGATATATTTTTCGAATAGCAGTGA	1305
QY	476	ThrLeuLysTyrSerAlaGlyGlySerLeu---SerAsnTyrGlnAsnThrThrPhePhe	494
DB	1306	GTATTTTCGATATTCATCCAAATTCAGAAATAATTAATAAAGAACTGATCTTAT	1365
QY	495	GlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPhe	514
DB	1366	ATGATTTCCAAACAAACATCGAAAT-----	1392
QY	515	AsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGly	534
DB	1393	AAAGAAATATGTCATCTATCTATATAAATAAAGCTGATAATATATATTTTTCAGTAGTT	1452
QY	535	LeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArg	554
DB	1453	AGAGAAAGAAAGAGAGTT-----GCATTTAGTTGGACACATACATAGTGTGATTC	1503
QY	555	TyrAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnLeu	574
DB	1504	CAAAATCAATAGATTTAGATAACATCCCAATTCAGCTCTAAAGCTTTGAAGGTA	1563
QY	575	AspThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeu	594
DB	1564	AGTTCTGATCGAAATTTGTGAAGGCTCTGTGCACAGGTGGAGACTTGTGTAATCTTT	1623
QY	595	GlnSerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPhe	614
DB	1624	AAAGATAGTAGGATTTTAGAGTTAGATTT---TTAAAAAATGTTTTCGACAATATCAA	1680

QY 615 IleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThr 634
Db 1681 GTACGTTATCGTTACTAATGCT- - - - - CCAAGACAAACAGTATTC 1725
QY 635 IleProGlyValIleGlyIleProGlnArgLeuAsnThrPheSerGlyThrAsn 654
Db 1726 TTAACCGGAATAGATACTATAAGTGGAG- - - CTCCCTAGTACCCTCCGCCAATAC 1782
QY 655 TyrAsn- - - - - AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThrVal 672
Db 1783 CCAAAATGCTACAGATTTAACATATGCAATGATTTGGATATGATCAATTTTCCAAAGAACAGTT 1842
QY 673 - - - - - ThrLeuProLeuAsnArgAsnIleProPheIle 683
Db 1843 CCAAAATAAACATTTGAAGGAGAGACACTTTTAAATGACC- - - - - TTA 1887
QY 684 PheAsnArgAlaAspValSerAsnSerIleLeuIleAspLysIleGluPheIlePro 703
Db 1888 TATGGTACCAAAATCATTTATATATATA- - - TATATTCACAAATCGAATTTATTCCTCA 1944
QY 704 IleThrSerSerMetHisGlnAsnArgGluLysGlnIleLeuGluThrIleGlnThrLys 723
Db 1945 ATCACTCAATCTGTATTAGATTATACAGAGAGCAAAATATAGAAAAACACAGAAATA 2004
QY 724 IleAsnThrPheThrAsnHisThrLysThrLeu 735
Db 2005 GTGAATGATTATTCTGTTAATTAACAAAGTTCTT 2040

RESULT 12
AAQ81180
ID AAQ81180 standard; DNA; 2061 BP.
AC AAQ81180;
XX
DT 25-MAR-2003 (revised)
DT 12-AUG-1995 (first entry)
XX
DE B.t. toxin PS71M3 gene.
XX
KW Delta-endotoxin; crystal protein; biological control agent; Calliphorid;
KW screw-worm; sheep blowfly; Lucilia; Phormia; Calliphora; insecticide;
KW pesticide; B.t.; ss.
XX
OS Bacillus thuringiensis.
XX
PN WO9502694-A2.
XX
PD 26-JAN-1995.
XX
PF 13-JUL-1994; 94WO-US007902.
XX
PR 15-JUL-1993; 93US-00093199.
XX
PA (MYCO) MYCOGEN CORP.
XX
PI Hickie LA, Payne J;
XX
XX
DR WPI; 1995-067338/09.
DR P-PSDB; AAR63079.
XX
PT Method for controlling Calliphoridae pests - specifically utilises
PT Bacillus thuringiensis isolates or toxins.
XX
PS Disclosure; Page 42-43; 50pp; English.
XX
CC A library was constructed from Bacillus thuringiensis PS71M3 total
CC cellular DNA in lambda Gem-11. Plasmid pMYC1636, selected in Escherichia
CC coli, contained a 15 kb insert expressing a beta-endotoxin gene. This was
CC sequenced (AAQ81180). A cured, acrystalliferous B.t. host carrying
CC pMYC1636 produced a 77 kDa crystal protein (AAR63079). (Updated on 25-MAR
CC -2003 to correct PN field.)
XX
XX

SQ Sequence 2061 BP; 779 A; 331 C; 319 G; 632 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3.86e-71 Length: 2061
Score: 899.50 Matches: 247
Percent Similarity: 48.32% Conservative: 126
Best Local Similarity: 31.99% Mismatches: 270
Query Match: 23.25% Indels: 129
DB: 2 Gaps: 30
US-10-783-417-2 (1-735) x AAQ81180 (1-2061)
QY 1 MetAsnGlnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 1 ATGAATCCATATCAAAATGAATGAATGAATATTCATATGCTCCATCAATGGTTT 60
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 61 - - - - - AGCAAGTCTAATACTATTCTAGATATCCATAGCAATAGCCAAATCAACA 114
QY 41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
Db 115 CTGAAACACGAAATTACAAAGATTGCTCAATGTGTCAAGATAATCAACAATATGGC 174
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIle 80
Db 175 AATAATCGGGGAATTTTGTAGTTCTGAAACTATTGTGAGATTAGTCAGGTATTATT 234
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGly 100
Db 235 GTAGTAGGAATATGTTAGGA- - - - - GCTTTGCTGCCCT- - - - - 270
QY 101 IleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120
Db 271 GTCTTAGCTGAGGTATAATATCTTTGGGACTTTGTTGCCGACTTTTGG- - - CAAGA 327
QY 121 GluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db 328 TCTGACCTCGAAATGTTGGCAGGATTTGTTAAACATCGGA- - - - - GGAAGG 375
QY 141 ProLeuThrGlu- - - - - SerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPhe 157
Db 376 CCTATACAGAAATAGATAAAACATAATTAATGTACTAATCTTCTATCGTAACACCTATA 435
QY 158 ArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeu 177
Db 436 AAAAACTCACTTGATAAATATCAAGAATTTTTCGATAAATGGAGCCAGCAGT- - - - - 489
QY 178 GlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIle 197
Db 490 - - - - - ACACAGCTAATGCTAAAGCAGTACATGATCTCTTTACT 528
QY 198 ArgPheGluAsnValHisAsn- - - - - AspPheIleArgGluIleProGlyPheGln 214
Db 529 ACCTTAGAACCTATAATAGATAAAGATTTAGATATGTTAAATAATAATGCTAGTATCGA 588
QY 215 LeuGluThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeu 234
Db 589 ATACCAACA- - - - - CTCCCTGCATATGCACAAATAGCTACTTGGCATTG 633
QY 235 AsnLeuLeuGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSer 254
Db 634 AATTTATTAACACATGCTGTCTACCTATTACAAATATATG- - - - - CTGCAAAATCAA 684
QY 255 GlnIleGluProAsnAlaGlyThrSerAspAspTyrTyrLys- - - - - LeuLeuLysGluAsn 273
Db 685 GGTATAAATCAAGTACTTTCAATTCATTAATTAATCTATCAGGGCTATTATAAAGCTAAA 744
QY 274 IleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArg 293
Db 745 ATACAAAGATATCTGACTATTGTATACAAACGTAACATGCAATGCAAGCACTATGATTAGA 804
QY 294 AspGluProAsnMetLysTrpSerIlePheAsnAspTyrArgTyrMetThrIleThr 313

```

Db      805 ACTAATACTAAACGCAATCGAATATCTATAATACTTACCGTTTAGAAATGACTCTAACT 864
Qy      314 valLeuaspThrIleSerGlnPheSerLeuTyrAspIleYsrArgTyrArgAspSerIle 333
Db      865 GTCTTAGATCTTATGCTATTTTCCAAATATTATGACCCAGAAATAAT- 912
Qy      334 GlyGlyIleGluVallySGlyIleLeuAsnGluLeuThrArgGluIleTyrThrGlu 353
Db      913 -----CCATAGAGAGTTAAATCTGAACCTTACGAGAGAGTTAT---ACGAAT 957
Qy      354 IleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyr 373
Db      958 GTTAATTTCAGATCAITTT-----AGAACCATAAACAGAACTAGAAAAT 999
Qy      374 AsnLeuThrArgAlaSerPheLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
Db      1000 GGATTAACTAGA---AATCCTACATTTATTCTTGGATAAACCAAGCGCGTTTTTACACA 1056
Qy      394 GluAsnThrAsn-----PheGlyAsnArgLeu 402
Db      1057 AGAAATTTCTCGAGACATCTTGTATCTTATGATATTTTCTTTTACAGGTAACCAAGATG 1116
Qy      403 ValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyr 422
Db      1117 -----GCCTTTACACATCTAATGATGATGATCCACATTAATCTGG 1155
Qy      423 GlyGluArgThrGly-----SerProThrThrIlysthrIleArgProPheGluSer 439
Db      1156 GGAAGCGGTTTCATCGACATATTTTCTCAAGACACATCCAAAGTATTTCTTTTATAGA 1215
Qy      440 Tyr-----LysValSerIleValThrAspArgGlnSerProProValSerPro 455
Db      1216 AACAAACCTATTGATAAGGTCGAAATTTGTACAGATGAGAGTACTCAGATATAATA- 1272
Qy      456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn 475
Db      1273 -----TATGAATGATATTTTTTTCGAATAGCAGTGAA 1305
Qy      476 ThrLeuLysTyrSerAlaGlyGlySerLeu---SerAsnTyrGlnAsnThrThrPhePhe 494
Db      1306 GTATTTTCGATATTCATCCATTCACACATAGAAAATAATTTATAAAGAACTGATCTTAT 1365
Qy      495 GlnPheProArgLysIlyAspCysAsnLeuValIleAspProGlyCysSerProAsnPhe 514
Db      1366 ATGATTTCCAAAACAAACATCGAAAAT----- 1392
Qy      515 AsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGly 534
Db      1393 AAAGAATATGGTCATCTCTATCGTATATAAATACTGATAATATATTTTTCAGTAGTT 1452
Qy      535 LeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArg 554
Db      1453 AGAAGAAGAGAGAGT-----GCATTTAGTTGGACACATACATAGTTGATTC 1503
Qy      555 TyrAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeu 574
Db      1504 CAAAATACAAATAGATTAGATAACATCACCCAAATCCACGCTCTAAAGCTTTGAAGGTA 1563
Qy      575 AspThrAsnSerLysValIleGluGlyProGlyHisThrGlyAsnLeuValTyrLeu 594
Db      1564 AGTCTCTGATTCGAAAATTTGTGAAGGTCCTGGTCACACAGGTGAGACTTGGTAATTCCT 1623
Qy      595 GlnSerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPhe 614
Db      1624 AAAGATAGTATGATTTTAGAGTTAGATTT---TTAAATAATGTTTCTCGACATAATCA 1680
Qy      615 IleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThr 634
Db      1681 GTACGTATTCGTTATGCTACTAATGCT-----CCAAAGACAAACAGTATTC 1725
Qy      635 IleProGlyValIleGlyIleProGlnArgLeuAsnAsnThrPheSerGlyThrAsn 654

```

1726 TTAACCGGAATAGATACTATAAGTGTGGAG---CTCCCTAGTAGCACCTTCCGCCAAAAC 1782

655 TyrAsn-----AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThrVal 672

1783 CCAAAATGCTACAGATTTTAAACATATGCAGATTTTGGATATGTAAACATTTCCAAAGACAGTT 1842

673 -----ThrLeuProLeuAsnArgAsnIleProPheIle 683

1843 CCAAAATAAACATTTGAAGGAGAGACACCTTTATTAAATGACC-----TTA 1887

684 PheAsnArgAlaAspValSerAsnSerIleLeuIleAspLysIleGluPheIlePro 703

1888 TATGGTACACCAATCATTCATATAATA---TATATTGACAAAATCGAATTTATTCCA 1944

704 IleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLys 723

1945 ATCACTCAATCTGTATTAGATTATACAGAGAAGCAAAATATAGAAAAAACACAGAAAAATA 2004

724 IleAsnThrPhePheThrAsnHisThrLysThrLeu 735

2005 GTGAATGATTTATTGTTTAAATTAATAAACAAAGTTCTT 2040

RESULT 13

AAN50525

ID AAN50525 standard; DNA; 3756 BP.

XX AC AAN50525;

XX XX 24-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 23-OCT-1991 (first entry)

XX XX Bacillus thuringiensis var. israelensis endotoxin insert in plasmid

DE DE PSY367.

XX XX Endotoxin; insecticide; ss.

XX XX Bacillus thuringiensis serovar israelensis.

XX XX EPI53166-A.

XX XX 28-AUG-1985.

XX XX 15-FEB-1985; 85EP-00301017.

XX XX 22-FEB-1984; 84US-00582506.

XX XX 22-JAN-1985; 85US-00693556.

XX XX (SYTR) SYNTRO CORP.

XX XX Walfield AM, Pollock TJ;

XX XX WPI; 1985-211724/35.

XX XX Polypeptide active against Diptera insects - prepd. from DNA sequence

PT PT coding for BTI endotoxin using bacterial host.

XX XX

XX XX Disclosure; Page 15a-e; 27pp; English.

XX XX The B. thuringiensis var. israelensis endotoxin insert in PSY367 is

CC CC expressed in a bacterial host. The protein produced has insecticidal

CC CC activity against dipteran insects. (Updated on 25-MAR-2003 to correct PA

CC CC field.) (Updated on 24-OCT-2003 to standardise OS field)

XX XX

SQ Sequence 3756 BP; 1379 A; 566 C; 659 G; 1150 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.: 4.2e-69 Length: 3756

Score: 881.00 Matches: 248

Percent Similarity: 47.42% Conservative: 120

Best Local Similarity: 31.96% Mismatches: 272

Query Match: 22.77% Indels: 136

DB: 1 Gaps: 31

QY 684 -----PheAsnArgAlaAspValSerAsnSerIleLeuIleAspIle 699
 Db 2819 AATGACCTTATTATGTGTACACCAATCATCATATATA---TATATTGACAAAT 2875
 QY 700 GluPheIleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThr 719
 Db 2876 GAATTTATTCCTCAATCACTCAATCTGTATTAGATTATACAGAGAAGCAATATAGAAAA 2935
 QY 720 IleGlnThrLysIleAsnThrPhePheThrAsnHisThrLysThrLeu 735
 Db 2936 ACACAGAAATAGTGTGATTTTATTGTTAAATTAACCAAGTTCTT 2983

RESULT 14

AAD43974

ID AAD43974 standard; DNA; 4896 BP.

XX AAD43974;

AC AAD43974;

XX AAD43974;

DT 13-DEC-2002 (first entry)

XX Bacillus thuringiensis ssp. finitimus cry28Aal gene.

XX Delta-endotoxin; cry26Aal; cry28Aal; insect-resistant plant; toxin;

KW transgenic host cell; insecticide; gene; ds.

XX Bacillus thuringiensis.

XX Key Location/Qualifiers

FH 1129.4458

CDS /tag= a

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

FT /product= "cry28Aal gene"

CC insect susceptible areas or plants to confer protection or resistance
 CC against harmful insects. The present sequence is *Bacillus thuringiensis*
 CC ssp. *finitimus* cry28Aal gene
 XX
 SQ Sequence 4896 BP; 1727 A; 757 C; 887 G; 1525 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4.8e-66 Length: 4896
 Score: 849.00 Matches: 229
 Percent Similarity: 48.76% Conservative: 106
 Best Local Similarity: 33.33% Mismatches: 248
 Query Match: 21.94% Indels: 104
 DB: 6 Gaps: 25
 US-10-783-417-2 (1-735) x AAD43974 (1-4896)
 QY 68 SerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAla 87
 Db 1210 AGCAGTGATACAGTCGCTGTAGTAGCGGAGGATGTAGTTGGGTACTATCTGACA 1269
 QY 88 GlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIle 107
 Db 1270 -----GCCTTTGCATCATTTGTTAATCCA-----GGTGGGTACTTATA 1308
 QY 108 SerPheGlyThrLeuIleThrValPheThrProAlaGlyGluGlnAsp---LysThrVal 126
 Db 1309 TCATTGGAACTTGGCTCCGCTTTCTTTGGCTGTATCAGAGGAGATCCAAAAAAT 1368
 QY 127 TrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeuThrGluSerIle 146
 Db 1369 TGGTCACATTTATGAAACACGAGAGACCTTTTAAATCAACAATTTCTACAGCTGTA 1428
 QY 147 LysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGlnSerTyrAsnThr 166
 Db 1429 AAGAAATAGCATTAGCTCATCTAAATGCTTTTAAAGATGATTATTAACGCTATGAAAGA 1488
 QY 167 AlaLeuAspAspThrArgLysLeuLysArgLeuGlnAlaProGlyLeuProSerSer 186
 Db 1489 GCATTTAATGATTGGAAGAGA-----AATCCAAGTGCA 1521
 QY 187 AlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPhe 206
 Db 1522 -----AATCTGCCAGATTGGTATCAGAGATTGTGAAACGCTCATTTCAATTT 1572
 QY 207 IleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeuLeuProIleTyr 226
 Db 1573 GTAAGCAATATGCCCAACTCCACTCCAGTATGACACATATTATTAAGTTCCTAT 1632
 QY 227 AlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGlyAlaGluLeuAlaAspGlu 246
 Db 1633 ACAGAAGCTGCNAATTTACATTTGAAATTTATATCATCAAGGTGTACAAATTCGGGATCAA 1692
 QY 247 TrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTyr 266
 Db 1693 TGGAAATCAGATCAACCAATTCACCAATGTTGAAGTATCATCAGGTACT-----TAT 1743
 QY 267 TyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyrArg 286
 Db 1744 TATCAGCAGCTATTGGTATATATTGAAAGATATATTATTTATTCACCAAGCATACCAT 1803
 QY 287 ThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTyr 306
 Db 1804 AAGAGATTGAATCACTTAAAGAAATCAGAAAAATCATATGGGATGCTTATAACACATAT 1863
 QY 307 ArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIle 326
 Db 1864 CGTCGAAATAGACCTTAATTTGATGATCTTGTCCCAACTTTTCTTTTATGATATA 1923
 QY 327 LysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThr 346
 Db 1924 CGTCGTTTT-----CCAGAGAGAGTAGAAGTAAATTAACA 1959
 QY 347 ArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeuArgValGlnPro 366

Novel isolated delta-endotoxin nucleic acid molecules, cry26Aal and cry28Aal isolated from *Bacillus thuringiensis finitimus*, that encodes toxin active against insects, useful for controlling insects.

Claim 1; Page 33-37; 42pp; English.

The invention relates to isolated delta-endotoxin nucleic acid molecules, cry26Aal and cry28Aal isolated from *Bacillus thuringiensis finitimus*, that encode a toxin that is active against insects. The invention is useful for producing an insect-resistant plant, by introducing the nucleic acid molecule into the plant, where the nucleic acid is expressible in the plant in an effective amount to control an insect. The invention is useful for producing a toxin that is active against insects by obtaining the transgenic host cell and expressing the nucleic acid molecule in the host cell, which results in the toxin that is active against insects. The toxin is useful for controlling an insect by delivering to the insect an effective amount of toxin. The invention is useful for controlling insects. The toxin is useful for inhibiting the ability of insect pest to survive, grow or reproduce, for limiting insect related damage or loss in crop plants, and to prophylactically treat

```

Db      1960 AGAGAGGTTTATACAAAGTTTAGATCATTTTAACACGACCCACCA----- 2001
QY      367  AsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeu 386
Db      2002 -----GGGCTATTACTTGGCTG 2019
QY      387  GluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIleSer 406
Db      2020 TCAGATATTGAGTTATACACGAGAGGTGGCAGAGCGGATTTATTCAGGTATT-- 2076
QY      407  AsnArgAspAlaProThrTyrSerAsnThr-----IleThrGluThrLeuTyrGly 423
Db      2077 ---CGAGAGCTCTAAATATTATATCTGCTGTAATCAATTTTACGATGATAAATAATTATGTT 2133
QY      424  GluArgThrGly---SerProThrThrLysThrIleArgProPheGluSerTyrLysVal 442
Db      2134 AATACAAATAGATTAAGTAAGCGCTCATTTACATTTATACCAGCGAA---TTATGACT 2190
QY      443  SerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIleIle 462
Db      2191 CACTTAAGCATAAACCGTCTCTTTTCAACAATAGCTGGTATAATAAGTTATACAGTTTA 2250
QY      463  AsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsnThrLeuLysTyrSerAlaGly 482
Db      2251 ATTCAAAAAATCGTATTCTCAACTTTTAAAAACGATAAT-----GAATATCAAAAAAAT 2304
QY      483  GlySerLeuSerAsnTyr-----GlnAsnThrThrPhePheGlnPheProArgLys 499
Db      2305 TTTAATGTGAATAATCAAAATGAACCTCAAGAACTACA----- 2343
QY      500  LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHis 519
Db      2344 -----AATCTCTAATGATTATGTGGTTCA---AACAGCCAAAAAATTCAAACAT 2391
QY      520  IleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIle 539
Db      2392 AATTATCTCTCAATTTCCATTAATCATCCAC-----AAGTTAGAGTTT 2433
QY      540  LeuAsp-----ThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyr 555
Db      2434 GCTGAGTATTTTCACTCTATATTTGCAATGAGTGGACACACAATAGTTAAACTCCCAA 2493
QY      556  AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp 575
Db      2494 AATTAAATATCAAGAGTGTGAGTACAAATCCCATTTGTTAAAGCTTACGAGTT--- 2550
QY      576  ThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln 595
Db      2551 ACTAACAATTTCAGTTATATAAGAGGACCGAGGTTTACAGGTGGAGATTAAATAGAACTTCGT 2610
QY      596  SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIle 615
Db      2611 GAT-----AAATGTTTCTATTAAATGTAAGCT---AGTCTTTTAAAAAATAACGTATA 2661
QY      616  ArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIle 635
Db      2662 AGCTCTATTTTATGCTCAATTAACGCAATAGCTGTATCAATAGCAGTAGTGATTCGGGA 2721
QY      636  ProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyr 655
Db      2722 GCAGGAGTTCTA-----TTGCAACTTACCTTTTCTAGAAAGGGAAC 2763
QY      656  AsnAsn-----LeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSer 670
Db      2764 AATAATTTTACAACTTCAAGACCTTAACTATAAGGATTTTCAATATCATACACTTTTAGTT 2823
QY      671  ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAsp----- 688
Db      2824 GATATTGAATTAATCCCGAAGAGTGAAGAAATTCATATCCATTTGAAGCGAGAGGATGATTAT 2883
QY      689  ValSerAsnSerIleLeuIleIleAspLysIleGluPheIleProIleThrSerSerMet 708

```

```

Db      2884 GAGGAGCGAGTGTCTTTTAATGTATAAATAGAGTTCAAACTATAGATGAAATATTAT 2943
QY      709  HisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePhe 728
Db      2944 -----ACTAATGAAATGAATTTTAGAGAGGCAAGAAAGCAGTGATGATTATTATTT 2994
QY      729  ThrAsnHisThrLysThrLeu 735
Db      2995 ATAACGCAACAACGCTTTG 3015

```

RESULT 15

```

ADF31301
ID      ADF31301 standard; DNA; 4896 BP.
XX      AC      ADF31301;
XX      DT      12-FEB-2004 (first entry)
XX      DE      Bacillus thuringiensis serovar finitimus pF2 DNA clone.
XX      KW      Cry26Aa1; cry28Aa1; delta-endotoxin; insect pest control;
KW      transgenic plant; insect resistance; insecticide; gene; ds.
XX      OS      Bacillus thuringiensis serovar finitimus.
XX      FH      Key      Location/Qualifiers
FT      CDS      1129..4458
FT      /tag= a
FT      /product= "Cry28Aa1 delta-endotoxin protein"
XX      US2003150018-A1.
XX      PD      07-AUG-2003.
XX      PF      15-JAN-2003; 2003US-00345020.
XX      PR      07-JAN-2000; 2000US-0175158P.
XX      PR      08-JAN-2001; 2001US-00756526.
XX      PA      (WOJC/) WOJCIECHOWSKA J A.
XX      PA      (LEWI/) LEWITIN E I.
XX      PA      (ZALU/) ZALUNIN I A.
XX      PA      (REVI/) REVINA L P.
XX      PA      (CHES/) CHESTUKHINA G G.
XX      PI      Wojciechowska JA, Lewitin EI, Zalunin IA, Revina LP;
PI      Chestukhina GG;
XX      DR      WPI; 2003-897623/82.
XX      DR      P-PSDB; ADF31302.
XX      PT      New isolated nucleic acid molecule encoding a toxin that is active
PT      against insects useful for controlling insect pests or for conferring
PT      insect resistance in plants.
XX      PS      Claim 7; SEQ ID NO 3; 42pp; English.
XX      CC      The present invention relates to the isolation of novel cry26Aa1 and
CC      cry28Aa1 delta-endotoxin genes from Bacillus thuringiensis ssp.
CC      finitimus. The sequences for the delta-endotoxin polypeptides are also
CC      disclosed. The invention provides methods for producing the toxins and
CC      compositions containing the toxins. The methods and sequences of the
CC      invention are useful for controlling insect pests in transgenic plants to
CC      confer insect resistance. The present sequence represents a DNA clone
CC      that contains the coding sequence for cry28Aa1 delta-endotoxin.
XX      SQ      Sequence 4896 BP; 1727 A; 757 C; 887 G; 1525 T; 0 U; 0 Other;

```

```

Alignment Scores:
Pred. No.:      4.8e-66      Length:      4896
Score:          849.00      Matches:      229
Percent Similarity: 48.76%      Conservative: 106
Best Local Similarity: 33.33%      Mismatches: 248

```

Query Match: DB:	21..94% 10	Indels: Gaps:	104 25
US-10-783-417-2 (1-735) x ADF31301 (1-4896)			
QY 68 SerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAla 87			
DB 1210 AGCAGTGATACAGTCGCTAGTAGTAAGCGCAGGAGTGTAGTGTGGGTACTATACGACA 1269			
QY 88 GlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIle 107			
DB 1270 -----CCCTTTGTCATCTTTGTTAATCCA-----GGTGTGGTACTTATA 1308			
QY 108 SerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGlnAsp---LysThrVal 126			
DB 1309 TCATTTGGAACTTGGCTCCGTTCTTTGGCTGATCCAGAGGAAGATCCAAAAAAT 1368			
QY 127 TrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeuThrGluSerIle 146			
DB 1369 TGGTCACAATTTATGAACACGAGAGACCTTTTAAATCAACAATTTCTACAGCTGTA 1428			
QY 147 LysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGlnSerTyrAsnThr 166			
DB 1429 AAAGAAATAGATAGCTCATCTAAATGGTTTAAAGATGTTAAGCTACTATGAAGA 1488			
QY 167 AlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeuProProSerSer 186			
DB 1489 GCATTTAATGATGAAGAGA-----AATCCAAGTGCA 1521			
QY 187 AlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPhe 206			
DB 1522 -----AATACTGCCAGATGGTATACAGAGATTTGAAACCGCTCATTTCAATTTT 1572			
QY 207 IleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeuLeuProIleTyr 226			
DB 1573 GTAAGCAATATGCCACAACCTCCAACTTCCACCGTATGCACATTTATTAAAGTTGCTAT 1632			
QY 227 AlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGlu 246			
DB 1633 ACAGAGCTGCAAAATTTACATTTGAATTTATATCAATCAAGGTGACAAATTCGCGGATCAA 1692			
QY 247 TrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTyr 266			
DB 1693 TGAAGTGCAGATCAACCAATTCACCAATTCACCAATTCACCAATTCACCAATTCACCAAT 1743			
QY 267 TyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyrArg 286			
DB 1744 TATGACGAGCTATTGGTATATATTGAAGAGTATATTATTATTTGCACCAACACATACCAT 1803			
QY 287 ThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTyr 306			
DB 1804 AAAGGATTTGAATCACCTTAAAGAAATCAGAAAAATACATGGGATGCTTATAACACATAT 1863			
QY 307 ArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIle 326			
DB 1864 CGTCGAGAAATACCTTAATTTGATTCGATCTTGTGCAACTTTCTCTTTTATGATATA 1923			
QY 327 LysArgTyrArgAspSerIleGlyIleGluValLysGlyIleLysAsnGluLeuThr 346			
DB 1924 CGTCGTTT-----CCAAGAGGATGAGAACTAGAAATTAACA 1959			
QY 347 ArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeuArgValGlnPro 366			
DB 1960 AGAGAGTTTATACAAGTTTATAGATCAATTTAAACAGCACCA-----2001			
QY 367 AsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeu 386			
DB 2002 -----GGCTATTACTTGGCTG 2019			
QY 387 GluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIleSer 406			
DB 2020 TCAGATATTGAGTTATACACGGAGGTGTCGAGAGCGGATTTATTATTCAGGTATT--- 2076			
QY 407 AsnArgAspAlaProThrTyrSerAsnThr-----IleThrGluThrLeuTyrGly 423			
DB 2077 ---CGAGAGTCTAAATATTACTGGTAATCAATTTTTCAGATGAAATAATATTATGTT 2133			
QY 424 GluArgThrGly---SerProThrThrLysThrIleArgProPheGluSerTyrLysVal 442			
DB 2134 AATACAAATAGATTAAAGTAAGCAGCTCATTTATTTACAGCGGAA---TTTATGACT 2190			
QY 443 SerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIleIle 462			
DB 2191 CACTTAAGCATAAACCGCTCTTTTCAACAATAGCTGGTATATAAATAGTTATACAGTTTA 2250			
QY 463 AsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsnThrLeuLysTyrSerAlaGly 482			
DB 2251 ATTCAAAAAATCGTATTCCAACTTTTAAACAGATAAT-----GAATATCAAAAAAT 2304			
QY 483 GlySerLeuSerAsnTyr-----GlnAsnThrThrPheGlnPheProArgLys 499			
DB 2305 TTTAATGTGAATAATCAAAATGAACCTCAAGAACTACA-----2343			
QY 500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHis 519			
DB 2344 -----AACTATCTTAATGATTTATGGTGTTC---AACAGCCAAAAATTTCAACAT 2391			
QY 520 IleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnGlnIle 539			
DB 2392 AATTATCTCTATTTTCCATTAATCATCCAC-----AAGTTAGAGTTT 2433			
QY 540 LeuAsp-----ThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyr 555			
DB 2434 GCTGAGTATTTTCACTCTATATTGTCATTTGGACACACAAATGTTAACTCCCA 2493			
QY 556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp 575			
DB 2494 AATTAAATATCAGAAAGTGTGAGTACACAAATCCCATTTGTTAAAGCTTTACGAAGTT--- 2550			
QY 576 ThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln 595			
DB 2551 ACTAAACATTCAGTTATTAAGAGAGCAGGTTTTCAGGTGGAGATTAAAGAACTCGT 2610			
QY 596 SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIle 615			
DB 2611 GAT-----AAATGTTCTTAATAATGTAAAGCT---AGTTCTTTAAAAAATACGCTATA 2661			
QY 616 ArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIle 635			
DB 2662 AGTCTATTTTATGCTGCAAAATACGCAATAGCTGTATCAATAGACGTAGGTATCCGGA 2721			
QY 636 ProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyr 655			
DB 2722 GCAGGAGTTCTA-----TTGCAACCTTACCTTTTCTAGAAAAAGGGAAC 2763			
QY 656 AsnAsn-----LeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSer 670			
DB 2764 AATAATTTTACAATTCAGACCTTAACATTAAGGATTTTCAATATCATACATTTTAGTT 2823			
QY 671 ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAsp----- 688			
DB 2824 GATATTGAAATACCGAAGTGAAGAAATTCATATCCATTTGAACGCGAGAGATGATTAT 2883			
QY 689 ValSerAsnSerIleLeuIleAspLysIleGluPheIleProIleThrSerSerMet 708			
DB 2884 GAGGAGGAGTGATTTCTTTTAAATGATAAATAGAGTTTCAAACTTATAGATGAAATAT 2943			
QY 709 HisGlnAsnArgGluLysGlnLeuThrIleGlnThrLysIleAsnThrPhePhe 728			
DB 2944 -----ACTAATGAATGAATTTAGAGAAGGCAAGAAAGCATGTAATGTTATTT 2994			
QY 729 ThrAsnHisThrLysThrLeu 735			
DB 2995 ATAAACGCAACAAACGCTTTG 3015			

Search completed: December 11, 2005, 23:38:24
Job time : 1173 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 11, 2005, 23:17:19 ; Search time 7267 Seconds

(without alignments)
4732.146 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869

Sequence: 1 MQQNNDNEVEIDSHSPY.....KLETIQTKINTFFNHTKTL 735

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 2339354128 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xlh
-Q=/cgn2_1/USPT.spool/US10783417/runat_01122005_141957_24358/app_query.fasta_1.903
-DB=EST -OFMT=fastap -SUFFIX=ret -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10783417 @CGN 1.1 5315 @runat_01122005_141957_24358 -NCPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	147.5	3.8	6567	11	DQ031803 Homo sapi
2	124	3.2	13107	11	DQ035722 Homo sapi
3	121.5	3.1	3477	10	CL977084 OsiFCC029
4	121	3.1	1754	10	AG392993 Mus muscu
5	121	3.1	2032	4	CNS08AVH
6	119.5	3.1	6777	10	CL967047 OsiFCC014
7	117.5	3.0	1986	10	AY416726 Mus muscu

8	117	3.0	2157	4	AK089255	8	AK089255	Mus muscu
9	115	3.0	1178	1	AJ538005	9	AJ538005	AJ538005
10	115	3.0	2322	10	CL960568	10	CL960568	OsiFCC004
11	115	3.0	2460	4	CNS09YP2	11	CNS09YP2	XB832541 Arabidops
12	114.5	3.0	820	7	CF951415	12	CF951415	UI-M-HL0-
13	112.5	2.9	1123	10	AG378595	13	AG378595	Mus muscu
14	112.5	2.9	2841	10	CL974457	14	CL974457	OsiFCC025
15	111	2.9	1143	4	CNS0A1JE	15	CNS0A1JE	XB829716 Arabidops
16	111	2.9	1616	10	AG396120	16	AG396120	Mus muscu
17	110.5	2.9	1105	3	BM463012	17	BM463012	AGN00012
18	109.5	2.8	752	7	CO096813	18	CO096813	GR_Ea20E
19	109.5	2.8	861	7	CO070923	19	CO070923	GR_Ea20E
20	109	2.8	1878	2	BF528790	20	BF528790	602041401
21	108.5	2.8	949	7	CV674212	21	CV674212	RET781_30
22	108.5	2.8	956	4	AY812233	22	AY812233	Schistoso
23	108.5	2.8	2282	4	AK044734	23	AK044734	Mus muscu
24	108.5	2.8	5001	4	CR860285	24	CR860285	Pongo pyg
25	107.5	2.8	2223	10	CL981417	25	CL981417	OsiFCC045
26	107.5	2.8	2439	10	CL982610	26	CL982610	OsiFCC048
27	107.5	2.8	2955	10	CL947058	27	CL947058	OsiFCC048
28	107	2.8	859	9	AZ548115	28	AZ548115	ENTDU91TR
29	106.5	2.8	1541	4	BC015498	29	BC015498	Homo sapi
30	106.5	2.8	1598	4	BC030976	30	BC030976	Homo sapi
31	106.5	2.8	2355	10	AY413036	31	AY413036	Mus muscu
32	106.5	2.8	2948	4	CNS0A4RI	32	CNS0A4RI	Arabidops
33	106.5	2.8	4751	4	CR859511	33	CR859511	Pongo pyg
34	106	2.7	881	9	BH162188	34	BH162188	ENTRY54TR
35	106	2.7	910	9	AZ546690	35	AZ546690	ENTEN48TF
36	106	2.7	2520	10	AG280133	36	AG280133	Mus muscu
37	106	2.7	3054	10	CL974410	37	CL974410	OsiFCC025
38	105.5	2.7	1092	10	CL065922	38	CL065922	CH216-105
39	105	2.7	622	7	CV176047	39	CV176047	SPHODROMA
40	105	2.7	697	7	CV533100	40	CV533100	LVS_035_B
41	105	2.7	747	3	BJ311235	41	BJ311235	BJ311235
42	105	2.7	16387	11	DQ038517	42	DQ038517	Pan trogl
43	105	2.7	16425	11	DQ038516	43	DQ038516	Homo sapi
44	104.5	2.7	458	1	AU006750	44	AU006750	AU006750
45	104.5	2.7	872	6	CD792427	45	CD792427	EST663788

ALIGNMENTS

RESULT 1	DQ031803	Homo sapiens	6567 bp	DNA	linear	GSS 02-JUN-2005
LOCUS	DQ031803	Homo sapiens	HC2093	gene, VIRTUAL TRANSCRIPT, partial	sequence,	
DEFINITION	DQ031803	genomic survey	sequence.			
ACCESSION	DQ031803	DQ031803.1	GI:86883012			
VERSION	DQ031803.1	GI:86883012				
KEYWORDS	GSS.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.					
AUTHORS	1 (bases 1 to 6567)					
TITLE	A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees					
JOURNAL	(er) PLOS Biol. 3 (6), E170 (2005)					
PUBMED	15869325					
REFERENCE	2 (bases 1 to 6567)					
AUTHORS	Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fedel-Alon,A., Tanenbaum,D.M., Civallo,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.					
TITLE	Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA					
JOURNAL						
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.					


```

FEATURES
  source      Location/Qualifiers
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
              /chromosome="12"
  gene        <1..>6567
              /locus_tag="HC2093"

ORIGIN
Alignment Scores:
Pred. No.:      0.000318      Length:      6567
Score:          147.50        Matches:      160
Percent Similarity: 33.63%    Conservative: 101
Best Local Similarity: 20.62% Mismatches:      238
Query Match:     3.81%       Indels:         277
DB:              11         Gaps:          40

US-10-783-417-2 (1-735) x DQ031803 (1-6567)
QY 10 TyrGluIleIleAspSerHisThrSerPro---TyrPheProAsnArgAsnSerAsnAsp 28
Db 1951 TATGAAGTGTCTATAAAATATAGATACCTTTATATATAGAAACACATCAACACAGAC 2010
QY 29 -----SerArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsnThr 44
Db 2011 ATAATATTAAGGAACCTTAAGACCTCACAC-----CTCTATAACATT 2052
QY 45 AsnTyrLysGluTrpLeuAsnMetCysGlnGlyAsnThrGlnTyrGlyAspAsnPheGlu 64
Db 2053 TCTGTAAGTCTTACACACAGATTGGTTCATGGCAATCAGGTA-----TCTTCTTTACTC 2106
QY 65 ThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSerGlyThr 84
Db 2107 TCTGTAAGGACTTCGGAGACTGTCCTGTATAGTGCACGAAATATACCTTACAAAAAT 2166
QY 85 LeuLeuAlaGlyIleGlyLeuThrSerIle-----SerGlyProIleGlyIleIle 102
Db 2167 ATTTCCTCTGGAGAGATTGAGCTATCATCTCTCCCAAGTAGTCCCAATGGAATCATA 2226
QY 103 GlyAlaIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGln 122
Db 2227 AAAAAATATACAAATTTTCTCAAGAGA-----AGTAATGGAATGAG 2268
QY 123 AspLysThrValTrpThrGlnPheIleLysMetGlyLullePheValAspThrProLeu 142
Db 2269 GAAAGAACTATAAATACA-----ACCTCTTTA 2295
QY 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeu-----154
Db 2296 ACCCAAAACATTAAAGTACTGAGAGAAATATACCAATATATCATTTAGGTGTCGTAGT 2355
QY 155 -----GluGlyPheArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAsp 170
Db 2356 ACACCTGAAAGGTGAAGGAGTTCGGAGTGCCTCCATAAGTATCTGACGGAGGAGAT--- 2412
QY 171 TrpArgLysLeuLysArgLeuGlnAlaProGlyLeuProSer-----SerAlaLeu 188
Db 2413 -----GTCCTCGATTCTCCCTCAAGACTTCTCTGTAAAA 2448
QY 189 GlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPheIleArg 208
Db 2449 CAGTCTCTGGTGTACGGTGAAGTGTGTATGTCATGGCAA-----2484
QY 209 GluIleProGlyPheGln-----LeuGluThrTyrLysThrLeuLeuLeuProIle 225
Db 2485 -----CCACCCCTGGAGCCAAATGGAATTTCTTTTATACAGTTTAT---GTC 2532
QY 226 TyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlyAlaGluLeuAlaAsp 245
Db 2533 TGGATAGATCATCATTAATAAACTATTAATGTCACTGAAACATCATTTGGAGTTATCAGAT 2592
QY 246 -----GluTrpAsnAlaAspIleHisProSer-----254

```

```

Db 2593 TTGATTATATAATGTTGAATACAGTCTTATGTAACAGCTAGCAGATTTTGGTGATGGG 2652
QY 255 -----GlnIleGluProAsnAlaGlyThr 262
Db 2653 AAAACAAGAACATATATCAATTAGCTTTCAAACACAGAGGAGCACCAGCGATCTCTCC 2712
QY 263 SerAspAspTyrTyrLysLeuLeuLysGluAsn-----273
Db 2713 AAAGATGTTTATTATGCAAACTCTAGTTCTTCATCAATAATTCTTTCTGGACACCTCT 2772
QY 274 -----IleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThr 287
Db 2773 TCAAAACCTAATGGGATTATACAATATTACTCTGTTTATTACAGAAATACTTTCAGTACT 2832
QY 288 GlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTyrArg 307
Db 2833 TTTATGCGAGNNNNNNNNNTCCAT-----GAAGTAAACCAATGACTTTGAC 2877
QY 308 ArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIle--- 326
Db 2878 AATATGACTGTATCCACAATTTATAGATAAACTGACAATATTTCAGCTACTATACATTTGG 2937
QY 327 LysArgTyrArgAspSerIleGlyGly-----IleGluVal---338
Db 2938 TTAACAGCAAGTACTTTCAGTTGGAAATGGAATAAAAGCAGTGACATCATTTGAAGTATAC 2997
QY 339 -----LysGlyIleLysAsnGluLeuThrArgGlu---IleTyrThr 351
Db 2998 ACAGATCAAGACATACCTCGAAGGGTTTGTGGAACCTGACTTACGAATCCATTTCGTCA 3057
QY 352 ThrGluIleAsnPheAspArgLeuProGlnLeuArg-----363
Db 3058 ACTGCATATAATGTAAGCTGGTCCACCGCTCAACCAAAAGGTCTAGTCTTCTACTAT 3117
QY 364 -----ValGlnProAsnLeuAlaThrMetGlu 372
Db 3118 GTTTCACCTGATCTTACAGCAGACTCTCTGCCATGTGTGAGACCCTCTTGTTCATATGAG 3177
QY 373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyr 392
Db 3178 AGAAGCATA-----TATTTTGTATAATCTGGAATAATACACTGATTAT 3219
QY 393 -----ThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIleSerAsnArg 408
Db 3220 ATATTAATAATTACTCTCATCAACAGAAAGGA-----3252
QY 409 AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySer 428
Db 3253 -----TTCTCTGATACCTATATCTGCCAGCTATAC-----3282
QY 429 ProThrThrLysThrIleArgProPheGluSerTyrLysValSerIleValThrAspArg 448
Db 3283 -----ATCAAGACTGAAGAA 3297
QY 449 GlnSerProValSerProIleGlnProHisPheIleIleAsnGlnIleGluLeuTyr 468
Db 3298 GATGTCACAGAACTTCACCAATA-----ATCAACTTTT 3333
QY 469 LeuAsnGlySerSerAsnAsnThrLeu-----LysTyrSerAlaGlyGly 483
Db 3334 AAAAACTTCTCTCTACTCTAGTTCTCTATCATGGGATCCCCAGTAGTAAGCAAAATGGT 3393
QY 484 SerLeuSerAsnTyrGlnAsnThrPhePheGlnPheProArgLysLysAspCysAsn 503
Db 3394 GCAATAATAAGTTATGATTTAACT-----TTACAAGGACCAAAATGAA---3435
QY 504 LeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIle-----520
Db 3436 -----AATTATCTTCTTCTACTTCTCTGAT 3459
QY 521 -----LeuSerHisPheSerLeuPheThrTyrSerTyrValIle 533

```

Db	3460	AATTACATAAATTGGGAAGAGCTTTCCACCAATTACATTATATAGCTTTTGGTCCGCGCA	3519
Qy	534	GlyLeuGlnLeuGlnIleLeuAspThrGlyValLeu---GlyTrpThrHisSerSerVal	552
Db	3520	AGAACTAGAAAAGGACTTGGTCCCTCCAGTATTCTTTTTCATACACAGATGAGTCAGTG	3579
Qy	553	-----AspArgTyrAsnAlaIleSerAspLysIle-----	562
Db	3580	CCGTTAGCACCTCCACAAAATTGACTTTTAATCAACTGTACTTCAGACTTTGTATGGCTG	3639
Qy	563	-----IleThrMetIleProAlaIleLysGlyAsn	572
Db	3640	AAATGGAGCCCAAGTCCTCTCCAGGTGGTATTGTTAAAGTATATAGTTTAAAAATTCAT	3699
Qy	573	AsnLeuAspThrAsnSer-----LysValIleGluGlyProGlyHisThrGlyGly	589
Db	3700	GAACATGAACTGACACTATATATTATAAGAAATATATACGGA---TTTAAAACTGAAGCC	3756
Qy	590	AsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIleThrCysGluThrProAsnSer	609
Db	3757	AAACTTGTGGACTGGAA-----CCAGTCAGC	3783
Qy	610	ThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn---GlyAlaGlyAsnThrLeu	628
Db	3784	ACC-----TACTCTACCGTGATCTCGGTTCACCAAGTTGGAATGGCAATCAATTT	3837
Qy	629	ProAsnIle-----SerLeuThrIleProGlyValIle	639
Db	3838	AGTAATGTAGTAAATTCACAACCCAGAAATCAGTTCAGATGTCGTG	3885
RESULT 2			
LOCUS	DQ035722	13107 bp DNA linear	GSS 02-JUN-2005
DEFINITION	Homo sapiens DNAH5 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.		
ACCESSION	DQ035722.1	GI:66886931	
VERSION	GSS.		
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 13107)		
JOURNAL	Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B., Hubisz, M.J., Fedel-Alon, A., Tanenbaum, D.M., Civeello, D., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.		
PUBLISHED	A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees		
REFERENCE	(e) PLoS Biol. 3 (6), E170 (2005)		
AUTHORS	15869325		
TITLE	2 (bases 1 to 13107)		
JOURNAL	Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B., Hubisz, M.J., Fedel-Alon, A., Tanenbaum, D.M., Civeello, D., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.		
COMMENT	Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA		
FEATURES	This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.		
source	Location/Qualifiers		
gene	1. 13107		
ORIGIN	/organism="Homo sapiens"		
Alignment Scores:	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	/chromosome="5"		
	<1. .>13107		
	/gene="DNAH5"		
	/locus_tag="HC4910"		

QY 345 -----LeuThrArgGluLeuTyrThrThrGluLeuAsnPhe 356
Db 4543 AGCTTTAAACCGTGGAGAGCTCTCTTGAGAGGAGACAGTACCTCGGAATCATCCGCC 4602
QY 357 AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThr 376
Db 4603 AACATGGAGACAGCTTGATGTTGCTGGGATCCCTACTGAGCAACAGGTACATATG--- 4659
QY 377 ArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheLeuPheTyrThrGluAsn--- 395
Db 4660 -----CCATTCAA-----GCCAGATTCAAAATGGGTGCGAGTACCTTTCCCACTCA 4707
QY 396 ThrAsnPheGlyAsnArgLeuValGlyIleSerAsn-----ArgAspAlaPro 411
Db 4708 ACAGACATCATCGAGAGCTGGATGAGCGTGCACAAACCTGTGGATTTATTAGAGCTGTC 4767
QY 412 ThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySerProThrThr 431
Db 4768 TTTGTGGGAGGAGACATTGCCAAGCAGCTGCCCAAGGAA-----GCC 4809
QY 432 LysThrIleArgProPheGluSerTyrLysValSerIleValThrAspArgGlnSerPro 451
Db 4810 AAGCGTTTCTTAACATAGATAAATCTTGGTGAAGATCATGACTCGGCACATGAAGTG 4869
QY 452 Pro-----ValSerProIleGlnProHisPhe 460
Db 4870 CCCAGTGTAGTCAGTGTGTGTGGAGATGAGACCTCGGGCAGCTGTGTACCACAC--- 4926
QY 461 IleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsnThrLeuLysTyrSer 480
Db 4927 TTGCTGGACAGTTGGAATA----- 4947
QY 481 AlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPheGlnPheProArgLysLys 500
Db 4948 TGCCAGAAATCCCTTACTGGGTACTTGGAGAAAACAGACTGTCTTCCTCGG----- 5001
QY 501 AspCysAsnLeuValIleAspProGly-----CysSerProAsn 513
Db 5002 ---TTTTTCTGCTCAGATCTGCGCTTCTAGAGATTCTGGGGCAGGGCTCGGACTCC 5058
QY 514 PheAsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIle 533
Db 5059 CACACTATACAGCCCATTTGCTGAATGTGTTGAC----- 5094
QY 534 GlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTrpThrHisSerValAsp 553
Db 5095 -----AACATTAATCTGTCAAG 5112
QY 554 ArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsn 573
Db 5113 TTCCAGAAAGATCTATGATCGAATCTGTCA--ATTTCCTCTCAAGAGGGTGAGACG 5169
QY 574 LeuAspThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeu---Val 592
Db 5170 ATTGAA-----TTGGATAAACCTGTCTATGCGAGAGGCAATGTGAAGTT 5214
QY 593 TyrLeuGlnSerGln--GlyArgLeuGluLeu-----ThrCysGluThrProAsn 609
Db 5215 TGGCTTAATTTCTTTTGAAGAATCTCAGTCTCATTTGCTGATTCGCCAGGCA 5274
QY 609 eThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuP 629
Db 5275 GCGGCAAA-----TATTCAAGAAACAGGTTTCCAACTAACTGAATTTCTTTCATCTTC 5328
QY 629 ro 629
Db 5329 CC 5330

RESULT 3

CL977084

LOCUS

DEFINITION

OsIFCC029775 Oryza sativa Expressed Library Oryza sativa (indica

cultivar-group) genomic, genomic survey sequence.

3477 bp DNA linear GSS 21-SEP-2004

GSS

GSS

GSS

ACCESSION CL977084

VERSION CL977084.1

KEYWORDS GI:52408678

SOURCE GSS

ORGANISM Oryza sativa (indica cultivar-group)

Oryza sativa (indica cultivar-group)

Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

1 (bases 1 to 3477)

Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,

Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,

Wong, G.K.S., Deng, X.W. and Wang, J.

An analysis of transcriptional regulation of the rice genome and

its comparison to Arabidopsis

Unpublished (2004)

Contact: Chen Chen

Department of Bioinformatic

Beijing Institute of Genomics

Chinese Academy of Sciences, Beijing 101300, China

Tel: 86-10-80481559

Fax: 86-10-80488676

Email: chenchen@genomics.org.cn

Rice genomic sequence.

Class: exon-trapped.

Location/Qualifiers

1..3477

/organism="Oryza sativa (indica cultivar-group)"

/mol_type="genomic DNA"

/db_xref="taxon:39946"

/clone_lib="Oryza sativa Expressed Library"

/notes="Oryza sativa exon trapped genomic sequences "

FEATURES

source

1..3477

/organism="Oryza sativa (indica cultivar-group)"

/mol_type="genomic DNA"

/db_xref="taxon:39946"

/clone_lib="Oryza sativa Expressed Library"

/notes="Oryza sativa exon trapped genomic sequences "

ORIGIN

Alignment Scores:

Pred. No.: 0.122

Score: 121.50

Percent Similarity: 31.81%

Best Local Similarity: 20.54%

Query Match: 3.14%

DB: 10

Length: 3477

Matches: 144

Conservative: 79

Mismatches: 243

Indels: 235

Gaps: 31

US-10-783-417-2 (1-735) x CL977084 (1-3477)

QY 132 LysMetGlyGluIlePheValAsp---ThrProLeuThrGluSerIleLys----- 147
Db 1105 AATATGAACAACTTACCTAGACCAAAACCAATACGGGTTCATCCCAAGAAATA 1164
QY 148 -----GlnLeuLysLeuGlnThrLeuGluGlyPheArg 158
Db 1165 GGAATGCTAGCAACCTGCAAGTTTTCAGCTAAGCAACACCTTGAGCGG----- 1218
QY 159 GlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGln 178
Db 1219 -----GAGATACCCACAGCTCTAGCAAAATCTGACCAACCTAGTACTTGA 1266
QY 179 AlaProGlyLeuProProSerSerAlaLeuGlnAlaLeuLeuLeu---LysIle 197
Db 1267 CTATATGTAATGACTGTGGGGCCCTATACCCAAAACTCTGCACCTCACCAGATG 1326
QY 198 ArgPheGluAsnVal---HisAsnAspPheIleArgGluIlePheGlnLeuGlu 216
Db 1327 CAATTACTTAGTCTTAGCAAAACAAACTTACTGCGAAATCCCTGCTGTTATCCAA 1386
QY 217 ThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeu 236
Db 1387 CTAAACAAAGTGGAAAAAATTTTACCTATACCAAAACCAA----- 1425
QY 237 LeuGlnGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIle 256
Db 1425 ----- 1425
QY 257 GluProAsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLys 276
Db ----- 276

```

1426 Db -----GTCACGGGTTCATCCAAA 1446
277 Qy TyrSerAsnTyrCysAlaAanThrTyrArgThrGlyLeuIysAsnLeuArgAspGluPro 296
1447 Db GAAATAGGAATGCTCCCAACCTGCAGCTTTTAGTCTAGGCAACAATACCTTGAACGGA 1506
297 Qy AsnMetIysTrpSerIlePheAsnAspTyrArgTyrMetThrIleThrValLeuAsp 316
1507 Db GAGATACCCACAACCTCTATCAAT-----CTAACCAACTAGAT 1545
317 Qy ThrIleSerGlnPheSerLeuTyrAspIleIysArgTyrArgAspSerIleGlyIle 336
1546 Db ACTTGT-----TCCCTATGGAT----- 1563
337 Qy GluValIysGlyIleIysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPhe 356
1564 Db -----AATGAACCTTCGGGCATATACCCCAAAACCTGACGCGTC 1605
357 Qy AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyr----- 373
1606 Db ACCAGATGCAATACCTTAGTCTTAGCAGCAACAACCTTACTGGCGAAATCCCTGGGTGT 1665
374 Qy -----AsnMetIysThrArgAlaSerPheLysLeuPheSerPheLeuGlu----- 387
1666 Db TTATCCCAATCTAACAAAATGGAA--AACTTTACCTATACCAAAACCAAGTCACGGT 1722
388 Qy -----GlnPheIlePheTyrThrGluAsn 395
1723 Db TCCATCCAAAAGAAATAGGAATGCTTCCCAACCTGCAGAGTTTTCAGCTAAGCAACAAC 1782
396 Qy ThrAsnPheGlyAsnArgLeuValGlyIleSerAsnArgAlaProThrTyrSerAsn 415
1783 Db ACCTTGACGGGGAGATATCCACAGCTCTATCAAT-----CTAACCAAC 1827
416 Qy ThrIleThrGluThrLeuTyrGlyGluArgThrGlySerProThrThrIleArg 435
1828 Db CTAGCTATTTCCTATGCTATGGGTAACTTCGGGGCTATACCCAA----- 1878
436 Qy ProPheGluSerTyrIysValSerIleValThrAspArgGlnSerProProValSerPro 455
1879 Db -----AAACTCTGATCTCACCAGATCCAAATATCTTGATCTTAGCAGT 1923
456 Qy IleGlnProHisPheIleIleAsnGln-----IleGluLeuTyrLeuAsnGlySer 472
1924 Db TCTCTTCTCGGAAATTCGAAATCTCACGGCATAGCAGACCTTTGGCTTGATAATAAC 1983
473 Qy Ser-----AsnAsnThrLeuIysTyrSerAlaGlyGlySerLeuSerAsnTyrGln 489
1984 Db TCATTTTCAGGACATTTGCTCGCAATGTATGATGGTGGGACACTCAAAACTTTTCATG 2043
490 Qy -----AsnThrThrPhePheGlnPheProArgIys---LysAspCys----- 502
2044 Db ATTGGTGGGAATGCCTTCGATGGCCCCATTCGGAAGCTTGAAGACATGTACAGCTTG 2103
503 Qy -----AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsn 516
2104 Db GTTAAACTTTCCGTTACATAAATCTGTTACGGGACACATATCTGAACATTTTGGAGTG 2163
517 Qy TyrSerHisIleLeuSer-----HisPheSerLeuPheThrTyr 529
2164 Db TATCCGCATCTCAAAATCGGTGAGTTCATACATAATAGGTTCTTTGGCGACATCTCACA 2223
530 Qy SerTyrValIleGlyLeuGlnLeuIleLeuAsp-----ThrGly 543
2224 Db AATGGTTCCTAGTCCCGAGCTTGAGAAATGATTTTCACAAAACATGATCAGAGT 2283
544 Qy ValLeuGlyTrpThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIle 563
2284 Db TTACTCAGACTTGATCAT-----AATAATCATGTGGTGAG----- 2319
564 Qy ThrMetIleProAlaIleLysGlyAsn-----AsnLeuAspThr 576
2320 Db -----ATTCAGCAGAAATTTGGAAATCTGAAAAGCCTATATAGATAAATTTATCATTC 2373
```

```

577 Qy AsnSerIysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSer 596
2374 Db AAT-----CAGCTCTCAGGATATTATACAGCACAGCTAGGAAGTTAAGCAAT 2421
597 Qy GlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616
2422 Db TTAGGTACCTTGATGATCAAGAAC----- 2448
617 Qy LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
2449 Db -----AATTGAGTGGACCAATATCCC 2469
637 Qy GlyValIleGly-----IleProProGln 644
2470 Db GATGAACCTAGGGGACTGCATCAGTTACAGGAAGTCTTCCATCATTCGTCAGCGCG 2529
645 Qy ArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGly 664
2530 Db ATGTGAGCTTAAACAGTCTTGTGTGTCTTCAACAACTTGGAA----- 2574
665 Qy TyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIlePro---PheIle 683
2575 Db ---GGACCTCTCCAGCAGGACATCTACTTCAA---AATGCATCAATAGCTGTTTAT 2628
684 Qy PheAsnArgAlaAspValSerAsnSer----- 692
2629 Db CACAATAAGTTATAGCATCTGCTCACCATAAGCCAAAGCTACTAAGTTTGTACTGCCA 2688
693 Qy -----IleLeuIleIleAspIysIleGluPheIleProIleThrSerSerMet 708
2689 Db ATTGTTCTTGTGTGGTCATTGTCTGCTGCTCAATATTATGTCATCAAAAGCTTGT 2748
709 Qy HisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePhe 728
2749 Db CATACAGAGAGAAACACACAAAGCTCTAGTGTATTACTGTGCAAGGAACATGTT 2808
729 Qy Thr 729
2809 Db TCT 2811

RESULT 4
AG392993 1754 bp DNA linear GSS 21-DEC-2004
LOCUS Mus musculus molossinus DNA, clone:MSMg01-211C07.TU, genomic survey
DEFINITION sequence.
ACCESSION AG392993
VERSION AG392993.1 GI:48018548
KEYWORDS Mus musculus molossinus (Japanese wild mouse)
SOURCE Mus musculus molossinus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Abe, K., Noguchi, H., Tagawa, K., Yuzuriha, M., Toyoda, A., Kojima, T.,
Ezawa, K., Saitou, N., Hattori, M., Sakaki, Y., Moriwaki, K. and
Shiroishi, T.
Contribution of Asian mouse subspecies Mus musculus molossinus to
genomic constitution of strain C57BL/6J, as defined by BAC-end
sequence-SNP analysis
Genome Res. 14 (12), 2439-2447 (2004)
15574823
2 (bases 1 to 1754)
Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
Direct Submission
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@psc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@tc.riken.jp).
```

Teukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Teukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY
Vector : pBACe3.6
R.Site 1 : ECoRI
R.Site 2 : ECoRI
Location/Qualifiers
1. .1754
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMG01-211007.TJ"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMG01 Mouse Male BAC Library"

FEATURES

source

ORIGIN

Alignment Scores:
Pred. No.: 0.0458 Length: 1754
Score: 121.00 Matches: 86
Percent Similarity: 38.59% Conservative: 51
Best Local Similarity: 24.23% Mismatches: 127
Query Match: 3.13% Indels: 93
DB: 10 Gaps: 16

US-10-783-417-2 (1-735) x AG392993 (1-1754)

```

Qy 413 TyrSerAsnThrIleThrGluThrLeuTyr-----GlyGluArgThrGlySerProThr 430
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1748 TACACATGCACATACAAATGAANCGTATACCTACTCAACACCGTACCATGCATCCANTC 1689

Qy 431 ThrLysThrIle-ArgProPheGluSerTyrLysValSerIleValThrAspArgGlnSe 450
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1688 ACA---ACGATANTCTCCCTATACACATCATATACATCAATCACACCTCATCAAGNA-- 1634

Qy 450 rProProValSerProIleGlnProHisPheIleIleAsnGlnIleGluLeu-----Ty 468
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1633 ----CCACATAAACACATACACACTCTATACTTCTCTATATACACAGACTCTCATCATCA 1578

Qy 468 rLeuAsnGlySerSerAsnAsnThrLeuLysTyrSerAlaGlyGlySerLeuSerAsnTy 488
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1577 TACACACATATCATCATCCTGCTGCTACA-----TCCATCCACAGATA 1539

Qy 488 rGlnAsn-----ThrThrPhePheGln-PheProArgLysLysAspC 502
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1538 TCACATCAACGACGACACACATCATCTGCTGCTCAAGATATATCCCATCAACAAATG 1479

Qy 502 ysAsnLeuVal-----IleAspProGlyCysSerProAsnPheAsnAsnTySer----- 518
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1478 CCTCTCTAATCAATCACTACTACAGTCACATGCTCACCACACACGTAACATATCATCAA 1419

Qy 519 -----HisIleLeuSerHisPheSerLeuPheThrTyrSerT 531
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1418 ACACACACAACTCAACTACATATATCTCCATACACATCATCAGCTC----- 1372

Qy 531 yrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerS 551
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1371 -----ACTCATCTCAC 1362

Qy 551 erValAspArgTyr-AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLys 570
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1361 ANTACAAATCATATCAATGCTACCTCTCTGTACACATACAAACATCTCCAACTCAACAA 1302

Qy 571 GlyAsnAsnLeuAspThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsn 590
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1301 TCAGATAACACACAAACAAATTCACCATCTCTATACATATACCCCAACACATGGCCATCA 1242

```

```

Qy 591 LeuValTyrLeuGlnSer-----GlnGly 598
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1241 CTCATCTCAATTACACAATACTTCAATATCAACAATCTTAAACAACCGTGGGGGT 1182

Qy 599 ArgLeuGlu-----IleThrCysGluThrProAsnSerThrGlnSer 612
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1181 CGAATCTCTCATCACACAACAAATCACACAACA-CAGAGCCCCCACTCAAAAAGCTCA 1123

Qy 613 TyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSer 632
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1122 TAT-----CTACATATAACACACATCTCTCTTAATTCATCTTTTAACGTAAC 1072

Qy 633 LeuThrIleProGlyValIleGlyIleProGlnArgLeuAsnAsnThrPheSerGly 652
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1071 ---ACATCTCCA-----ACCCCATCCCATATAAGACAAAATATTAGCCCC 1030

Qy 653 ThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThrVal 672
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1029 ACATCCT--AATCAA-----CACACTCATATTCCAAACCATAG 993

Qy 673 ThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSer 692
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 992 ACACCTCAATCATATAAACAATCCCAACATGACCACTAAATAACATATATATAC 933

Qy 693 IleLeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGln----- 710
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 932 ATATTATATCTCCCACTGCAACAACAATCCCAACATCACAATACCGGTACCATAAACC 873

Qy 711 AsnArgGluLysGlnLysLeuGluThrIleGlnThrLys 723
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 872 AACAAACGAAACAACAACAATACAAACCAATAATAACAAG 834

```

RESULT 5

```

CNS0A8VH 2032 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GS1TFB40ZE03 of Flowers and buds of strain col-0 of Arabidopsis
thaliana (thale cress).
ACCESSION BX819057.1 GI:42467943
VERSION BX819057.1
KEYWORDS HTC; GSUT; cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 2032)
AUTHORS Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2032)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

```

Db	1041	-----GGCCTTATATCCAGGTCTATGACCAACGCCCATATG--	1076
Qy	547	TrpThrHisSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIle	566
Db	1076	-----	1076
Qy	567	ProAlaIleLysGlyAsnAsn-----LeuAspThrAsnSerLysValIleGlu	582
Db	1077	-----TGCCAAGGTAAATGACGACGCCATATATGCCAGGTAATGACGACGCCCATAT	1130
Qy	583	GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle	602
Db	1131	ATCCAGGTCATGAGCAGCGGTATATATGCCAGCTCAATCTCAGTCTCAGTCTCAACCT	1190
Qy	603	ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaIleThrAsn	622
Db	1191	CAACCTCAA---CCTCAACTCAGCAACATATATGCTGCGACCGCAACCTCGAATCAAC	1247
Qy	623	GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIle-----	639
Db	1248	ATCAGGCGCCCTCTTCAACCTTAACCAAGTATCTACCAACCACCGCGGTAGTCTCTAGCCCA	1307
Qy	640	-----GlyIleProGlnArgLeuAsnAsnThrPheSerGlyThr-----	653
Db	1308	GTGCGCTATAATCTTCTCTCTCGGTCAATGCTGTCCTGACCTGTGACGCTACTCTCAG	1367
Qy	654	-----AsnTyrAsnAsnLeuGlnTyrGlyAspPhe	663
Db	1368	CCCTCTCCCATGTTCTAGTCAGTATGTTGGATTTCTTCTCTCGATATAACGGTTTT	1427
Qy	664	GlyTyrPheGlnPheProSerThrValThrLeu---ProLeuAsnArgAsnIleProPhe	682
Db	1428	GGACCACTGCGAGTCACCTACATCCCAATTCTTCAACCATCTCTCCTACCGGTTACCCGAAT	1487
Qy	683	IlePheAsn-ArgAlaAspValSerAsn-----SerIleLeuIleI	696
Db	1488	ATGTTCTCTCCGAGATCGCCTTACCCATTTCTGTCTACCGAGGTTCAGTATCTCTC	1542
Qy	696	eAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGluLysGlnIle	716
Db	1543	-----AACCATTACCCCAACTTCTCATTTTCAAAATTGCTCAACAGGAAGT	1592
Qy	716	sLeuGluThrIleGlnThrLysIleAsnThrPhePheThr	729
Db	1593	CTTGGACCGGTGTCAGTCCAAGTCAAGTCTCTCTCTCAGC	1632
RESULT 6			
LOCUS	CL967047	6777 bp	DNA linear
DEFINITION	OaIFC014664	Oryza sativa Expressed Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.	GSS 21-SEP-2004
ACCESSION	CL967047		
VERSION	CL967047.1	GI:52388742	
KEYWORDS	GSS.		
SOURCE	Oryza sativa (indica cultivar-group)		
ORGANISM	Oryza sativa (indica cultivar-group)		
REFERENCE			
AUTHORS	Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G.K.S., Deng, X.W. and Wang, J.		
TITLE	An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis		
JOURNAL	Unpublished (2004)		
COMMENT	Contact: Chen Chen Beijing Institute of Genomics Chinese Academy of Sciences, Beijing 101300, China Tel: 86-10-80481559 Fax: 86-10-80488676 Email: chenchen@genomics.org.cn		

Db	1993	GGGACTATG				1931
Qy	698	siLeGluPheilleProilleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuG1				718
Db	1932	-----GTGCCCCCAAGTTCTTCACAGANTAGAAAGAGAGAAATTTGA				1975
Qy	718	uThrIleGlnThrLysIleAsnThrPhePheThrAsnHisThr				732
Db	1976	CCAGATTGTTGATTCCCAATAGTTGGTTTCTTGTCACTCACAGT				2018
RESULT 7						
AY416726						
LOCUS	AY416726					
DEFINITION	Mus musculus LY64 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.					
ACCESSION	AY416726					
VERSION	AY416726.1	GI:39772686				
KEYWORDS	GSS.					
SOURCE	Mus musculus (house mouse)					
ORGANISM	Mus musculus					
	Eukaryota; Chordata; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Mus.					
REFERENCE	1 (bases 1 to 1986)					
AUTHORS	Clark,A.G., Gulanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.					
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios					
JOURNAL	Science 302 (5652), 1960-1963 (2003)					
PUBMED	14671302					
REFERENCE	2 (bases 1 to 1986)					
AUTHORS	Clark,A.G., Gulanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.					
TITLE	Direct Submission					
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA					
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.					
FEATURES	Location/Qualifiers					
source	1..1986					
gene	/organism="Mus musculus"					
	/mol_type="genomic DNA"					
	/db_xref="taxon:10090"					
	<1..>1986					
	/gene="LY64"					
	/locus_tag="HCM5986"					
ORIGIN						
Alignment Scores:						
Prod. No.:	0.143	Length:	1986			
Score:	117.50	Matches:	114			
Percent Similarity:	31.72%	Conservative:	76			
Best Local Similarity:	19.03%	Mismatches:	190			
Query Match:	3.04%	Indels:	221			
DB:	10	Gaps:	22			
US-10-783-417-2 (1-735) x AY416726 (1-1986)						
Qy	201	AsnValHisAsnAspPheIleArgGluIleProGly				212
Db	112	AACTGTGAATAATTTAGTGTCTCAATGAATTTCTGGCACTTTACCAACTCAACAGAAATGT				171
Qy	213	-----PheGlnLeuGluThrTyThrLysThrLeuLeuProIleTyAlaGlnAlaAa				230
Db	172	TTGGATTACCTTTAATGTCTTGGCTTACCATTCAAAACACGACCTTCAGCAGACTTATC				231
Qy	231	AsnPheHisIleLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGlu				246

QY 528 hrTy-SerTyrrValIleGlyLeuGlnLeuGlnIleLeuAspThrGly----- 543
Db 1238 CTGAGCATTCAGAAATCCCTCAGTACGACTCTAGATTTGGCATTTACTCGACTAA 1297
QY 544 -----ValLeuGlyTrpT 548
Db 1298 AGTGAAGATGCACAGAGTCCCTCCAGAACCTCCATCTTTGAAGGTGCTGAATCTCT 1357
QY 548 hrHisSerValAspArgTyrrAsnAlaIleSerAspIlySilleIleThrMetIleProA 568
Db 1358 CCCACAGCTCTTGAC-----ATCAGCAGTGCAGAGCTCTTCGATGGCTGCCAG 1408
QY 568 laile-----LysGlyAsnAsnLeuAspThrA 577
Db 1409 CACTCCAGCATTTGAACCTTACAGGGAATCACTTTCCAAAAGGAATATCCAAAAGACCA 1468
QY 577 snSerIysValIleGluGlyProGlyHisThrGlyAsnLeuValTyrrLeuGlnSerG 597
Db 1469 ACTCA-----CTTCAGACAC 1483
QY 597 lngIyArgLeuGluIleThr-----CysGluThrProAsnSerThrGlnSert 613
Db 1484 TGGGAGACTAGAAATCTGGTTTATCTCTTTTGATCTCTCTCCTCATTGACACGACG 1543
QY 613 yrPheIleArgLeuArgTyrrAlaThr----- 621
Db 1544 CCTTCACCATCTGAAGATGATGAATCATGTAGCCTGAGTCAACAAGCGCTGACATCCA 1603
QY 622 -----AnGlyAlaGlyAsnThrL 628
Db 1604 GTAGCATTCAGGCTCTTACTCATCTTAAAGGGATCTTACCTCAGTGGCTCCAAAT- 1659
QY 628 euProAsnIleSerLeuThrIleProGlyValIleGlyIleProProGlnArg 645
Db 1660 -----CACATCATCATCTTACCAGTCTCTCCCATCTTGTCCAGCAG 1707

RESULT 8
AK089255
LOCUS
DEFINITION
Mus musculus NOD-derived CD11c +ve dendritic cells cDNA, RIKEN full-length enriched library, clone:F630107B15 product:lymphocyte antigen 78, full insert sequence.
ACCESSION
AK089255
VERSION
AK089255.1 GI:26354375
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
Carninci, P. and Hayaishizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
10349636
REFERENCE
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayaishizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
11042159
REFERENCE
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Hara, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayaishizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL
PUBMED
REFERENCE
AUTHORS
THE RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
JOURNAL
REFERENCE
AUTHORS
THE FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
JOURNAL
REFERENCE
AUTHORS
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Hayano, M., Hanagaki, T., Hara, A., Hashizume, W., Hayaishida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, R., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akai, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayaishizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayaishizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/
FEATURES /
source
Location/Qualifiers
1. 2157
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM.DB:F630107B15"
/db_xref="taxon:10090"
/clone="F630107B15"
/cell_type="NOD-derived CD11c +ve dendritic cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
134. 2119
/note="unnamed protein product; lymphocyte antigen 78 (MGD|MGRI:1194924, GB|NM_008533, evidence: BLASTN, 99%, match=2136)
putative"
/codon_start=1
/protein_id="BAC40816.1"
/db_xref="GI:26354376"
/translation="MAPDISCFVLVALFLASCRATTSDDQKICKEVNTKNCENLGL NEIPGLTPNSTECLFESFNVLPTQNTFFSRNLNLTDFDLTRCQYIWHEDTFQSOHR LDTVLITANPLIFMAETLSPGKALHFLFIQTGISIDFIPLNKQTLSELYGNSH ISSIKLPKPTTEKLVLDLQFNNAHYLSKEDSSLOQATNLNLNGNDIAGIELGA FDSAVQSLNFGGTQNLVIFKGLKNSITQSLWLTGTFEDMDDEIDSPAVFGLCEMSV ESNLQKHLVFNISNTTFCFSGQLGDLTATHLSELPSGLVLTAKLVLSANKFE NLQCISASNPFSPLTHLSIKNTKRLKLEGTGCLENLENRLDLSDHDDTSDCCNLQL RNLSHQSLNLSYNELSLKTEAPKCEPQLELLDLATFLRKVKQAQSPFNHLHLKVL NLSSHLLDISSEQLFDGLPALQHLNLQGNHFPKGNIOKTNLSQTLGLRLEILVLSFCDL

SSIDQHAFTSLKMMHVDLSHNRILTSSIEALSHLKIYLYNLASNRISIIPLSLPLIL
SQORTNLRONPLDCTCSNIYFLEWYKENMKLEDTEDTLCENPPLLRGLRSLSDVTL
CSMAVGIIPLIIVFLVAILLIPAVKYFLRWKYQHI"

ORIGIN

Alignment Scores:

Pred. No.: 0.188 Length: 2157
Score: 117.00 Matches: 114
Percent Similarity: 31.72% Conservative: 76
Best Local Similarity: 19.03% Mismatches: 190
Query Match: 3.02% Indels: 221
DB: 4 Gaps: 22

US-10-783-417-2 (1-735) x AK089255 (1-2157)

QY 201 AnValHisAsnAspPheIleArgGluIleProGly----- 212
DB 245 AACTGTGAATAATTAGTCTCAATGAATAATCTCGGCACTTTACCAAACTCAACAGAAATGT 304
QY 213 -----PheGlnLeuGluThrTyLysThrLeuLeuLeuProIleTyAlaGlnAlaAa 230
DB 305 TTGAGGTTGAGCTTTAATGCTTCCTACCATCAAAACAGGCTTCAGAGACTTATC 364
QY 231 AnPheHisLeuAsnLeuGlnGlnGlyAlaGluLeuAlaAspGlu----- 246
DB 365 AATCT-CACCTTCTCGATTAAACAGGTCAGATTACTGGATACATGAAGATCTTT 423
QY 246 ----- 246
DB 424 CCAAGCCCAACATCGGTTAGACACCTTGTAACCGCAAAATCCCTGATTTATGGC 483
QY 247 -----TrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGly 261
DB 484 AGAGACAGCACTTAGTGGGCTAAGCAATTGAAGCATCTCTTCATCCAAACAGGAAT 543
QY 262 -ThrSerAspAspTyTyLysLeuLeuLysGluAsnIleProLysTySerAsnTyCy 281
DB 544 ATCCAGTATTGATTTATCCCACTGCACAAATCAAAACCTTGAA---AGTCTCTATCT 600
QY 281 salalaThrTyArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpse 301
DB 601 CGAAGCAACCAATATTTCCTCCATTAAGCTCCCAAGGTTTCCCAACAGAGAAGCTGAA 660
QY 301 rIlePhe-----AsnAspTyArgArgTyMetThrIleThrValLeuAspThrI 318
DB 661 GTCCTCGGATTTTCAGATAAATCTATCCATACCTGCTCTAAAGAGATATGAGCTCT 720
QY 318 eSerGlnPheSerLeuTyAspIleLysArgTyArgAspSerIleGlyGlyIleGluVa 338
DB 721 CCAGCAGCCCAATCTGAGCCTTAACCTTAATGGAATGACATTCGAGGAATAGAGCT 780
QY 338 1----- 338
DB 781 GGGGGCTTCGACTCAGCTGTCTCCAAAGTTTAAACTTTGGAGGACTCAGAACTTGCT 840
QY 339 -----LysGlyIleLysAsnGluLeuThrArgGluIleTyThrThrGluLeAs 355
DB 841 AGTTATCTTCAAGGGTTTGAAGAACTCTACGATCCAGTCTCTTTGGCTGGGACATTTGA 900
QY 355 n-----PheAspArgLeuProGlnLeuArgVa 364
DB 901 GGACATGGATGAGGAAGATATTAGTCTCGCTGGTGTGAGGGTCTCTGTGAATAATGTCTGT 960
QY 364 lGlnProAsnLeuAlaThrMetGluTy-----AsnLeuThrArgAlaSerPheLysLe 382
DB 961 GGAGAGCATCAACCTACAGAGCATATTATCTTCAACATTTCTCCAAACATCTCCATTC 1020
QY 382 uPheSerPheLeuGluGlnPhe---IlePheTyThrGluAsnThrAsnPheGlyAsnAr 401
DB 1021 CTTTCACTGGCTCCAGGAACCTGGACCTTAACAGCCACTCACCTGAGTGAATTCCTCTCG 1080
QY 401 gLeuValGlyIleSerAsnArgAspAlaProThrTySerAsnThrIleThrGluThrLe 421

Db 1081 ACTTGTGGGACTAAGC----- 1096
QY 421 uTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGlu-SerTyL 441
Db 1097 -----ACACTTAAGAAATTAGTTCTCAGTGC 1122
QY 441 ysValSerIle---ValThrAspArgGlnSerProProValSerProIleGlnProHisP 460
Db 1123 AATAAGTTTGAAGATTTGTGCCAAATCAGTGTTCCTCAATTTCCCTCC-CTTACTCAC 1181
QY 460 heIleIleAsnGlnIleGluLeuTyLeuAsnGlySerSerAsnAsnThrLeuLysTyS 480
Db 1182 TTTCCATC-----AAGGCCAACACGAGAGACTTG 1211
QY 480 erAlaGly---GlySerLeuSerAsnTyGlnAsnThrThrPhePheGlnPhePro---- 497
Db 1212 AGTCGGTACTGCTGTTTGAAGAAACCTAGAAATCTCCGGAACTTGACCTCAGCCATG 1271
QY 498 -----ArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheA 515
Db 1272 ATGACATTGAAACTTCTGATTGT-----TGCACCTGCAACTCC 1310
QY 515 snAsnTySerHisIleLeuSer-----HisPheSerLeuPheT 528
Db 1311 GAAACCTGTCTCACTTACAGAGCCTGAACCTGAGCTACAATCAACCCCTGAGCCTCAAGA 1370
QY 528 hrTySerTyValIleGlyLeuGlnIleLeuAspThrGly----- 543
Db 1371 CGGAGGCATTCAAAGAAATGCCCTCAGCTAGAACTCTAGATTGGCAATTACTCGACTAA 1430
QY 544 -----ValLeuGlyTpt 548
Db 1431 AGGTGAAGATGCACAGAGTCCCTTCCAGAACCTCCATCTTTTGAAGTGTGTAATCTCT 1490
QY 548 hrHisSerSerValAspArgTyAsnAlaIleSerAspLysIleIleThrMetIleProA 568
Db 1491 CCACAGCTCTCTGAC-----ATCAGCAGTGAAGCAGCTCTTCGATGGCTGCCAG 1541
QY 568 laIle-----LysGlyAsnAsnLeuAspThrA 577
Db 1542 CACTCCAGCATTTGAACCTTACAGGGAATCACTTTCCAAAGGGAATATTCAAAGACCA 1601
QY 577 snSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyLeuGlnSerG 597
Db 1602 ACTCA-----CTTCAGACAC 1616
QY 597 lnglyArgLeuGluIleThr-----CysGluThrProAsnSerThrGlnSerT 613
Db 1617 TGGGAGACTAGAAATCCTGTTTATCTCTTGTGATCTCTCTCCATTCGACGACGACG 1676
QY 613 yrPheIleArgLeuArgTyAlaThr----- 621
Db 1677 CTTTACCAGCTCTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1736
QY 622 -----AsnGlyAlaGlyAsnThrL 628
Db 1737 GTAGCATTTGAGGCTCTTAGTCTCTTAAGGGGATCTACTCAATCTGGCTCCCAAT--- 1792
QY 628 euProAsnIleSerLeuThrIleProGlyValIleGlyIleProGlnArg 645
Db 1793 -----CGCATCAGCATCATCTCTCCAGTCTCTCTCCCAATCTTGTGCCAGCAG 1840

RESULT 9

LOCUS AJ538005 1178 bp mRNA linear EST 01-JUL-2004
DEFINITION AJ538005 1A1 Timarcha balearica cDNA clone Timarcha6B7, mRNA
SEQUENCE
ACCESSION AJ538005
VERSION AJ538005.1 GI:40311007
KEYWORDS EST.
SOURCE Timarcha balearica
ORGANISM Timarcha balearica
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 Chrysomeloidea; Chrysomelidae; Chrysomelinae; Timarchini; Timarcha.
 1 (bases 1 to 1178)
 Gomez-Zurita, J., Kopliku, F., Theodorides, K. and Vogler, A.P.
 Resources for a phylogenomic approach in leaf beetle (Coleoptera)
 systematics
 (in) Jolivet, P., Santiago-Blay, J.A. and Schmitt, M. (Eds.); NEW
 DEVELOPMENTS IN THE BIOLOGY OF CHRYSOMELIDAE, 19-35 (2004)
 Contact: Gomez-Zurita J
 Department of Entomology
 The Natural History Museum
 Cromwell Road, London SW7 5BD, United Kingdom.
 Location/Qualifiers

FEATURES
 source
 1..1178
 /organism="Timarcha balearica"
 /mol_type="mRNA"
 /db_xref="taxon:79517"
 /clone="Timarcha687"
 /dev_stage="adult"
 /clone_lib="IAL"
 /note="country: Spain;Balearic Islands, Mallorca, Esporles"

ORIGIN
 Alignment Scores:
 Pred. No.: 0.12 Length: 1178
 Score: 115.00 Matches: 75
 Percent Similarity: 34.92% Conservatives: 35
 Best Local Similarity: 23.81% Mismatches: 106
 Query Match: 2.97% Indels: 99
 DB: 1 Gaps: 15

US-10-783-417-2 (1-735) x AJ538005 (1-1178)

QY 470 AsnGlySerSerAsnThrLeuLysTyrSerAlaGlyGly----- 483
 Db 227 AATGCTCAGCGAACAC-----AATTATCTCTTAATAATAATTATCAAGTAATGCT 280
 QY 484 -----SerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPhePro---Arg 498
 Db 281 AATTATTACCAAGTAACGGGAATCATCAAGGAACAATGGGACAGATTTCATCATCG 340
 QY 499 LysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnTyrSer 518
 Db 341 AATAACGATAGATATCAAAAGCTACACGGAATCCCTTTCCAAACATGAATAATAACGGG 400
 QY 519 HisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGln 538
 Db 400 ----- 400
 QY 539 IleLeuAspThrGlyValLeuGly-----TrpThrHisSerSerValAspArgTyrAsn 556
 Db 401 -----GACCAAGGAATATGGAATGCTATCCCATTCAGGAATAGTGATACAC 454
 QY 557 AlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnLeuAspThr 576
 Db 455 AATAATAATCCCAAGA-----AATAATCTCGCAGGAATAATTGGAATTTT 499
 QY 577 AsnSerLysValIleGlyProGlyHisThrGlyAsnLeuValTyrLeuGlnSer 596
 Db 500 AATTCCTGCACCAAGTAATGATATTCAGGAATAGTGTCATGCTTATCCATCACAAGAC 559
 QY 597 Gln-----GlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyr 613
 Db 560 AAAAATTATCCAGGAATAATTGGAATAGCTCTCTTCACCAAGAAAT---AATAGCTAT 616
 QY 614 PheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeu 633
 Db 617 -----ACAGGAATAGTGCAATGCTTATCCATCACAACAAAT 655
 QY 634 ThrIleProGly-ValIleGlyIle---ProProGlnArgLeuAsnAsnThrPheSerGly 652
 Db 656 AATTATCCAGGAATAATTGGAATATCTTCACCTTCACCGGAATAATACTCTTATCCAGGA 715

QY 653 ThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPhe----- 666
 Db 716 ACTAATGGGAATACTTACCAAAATCGGAATGGTGAATACTATCCAAAGAATAATGGTAAT 775
 QY 667 GlnPheProSerThrValThr----- 673
 Db 776 AATTTTCCTTCAGTGAATACCAAGTTCATATCCAGGAAGCAGTGGCAATAATTATCCAGGC 835
 QY 674 LeuProLeuAsnArgAsnIleProPheIlePheAsn-----ArgAlaAspValser 690
 Db 836 TTCAATATAATAATAATAATTATTATTCAGTGCATAACGCCAACGGGAATACTGATTATGCT 895
 QY 691 AsnSerIleLeuIleIleAspLysIle-----GluPheIleProIleThr 705
 Db 896 GGCAGTGTGGGAATACGACAGGAATGCGGGAACCTTTTCGGAAGTCAAAATCCCTAGCACT 955
 QY 706 -----SerSerMetHisGlnAsn----- 711
 Db 956 CAACCAAAAATTTTACTGGCATAGCGGTAGTGGGCCCACTTACATAAAAATAATATGGG 1015
 QY 712 -----ArgGluLysGlnLysLeuGluThrIleGlnThr 722
 Db 1016 ACGGCCAATGAAGAAGAGTGTGAGTAATGTTAAACG 1054

RESULT 10

CL960568

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT: Chen Chen

Department of Bioinformatic

Beijing Institute of Genomics

Chinese Academy of Sciences, Beijing 101300, China

Tel: 86-10-80481559

Fax: 86-10-80488676

Email: chenchen@genomics.org.cn

Rice genomic sequence.

Class: exon-trapped.

Location/Qualifiers

1..2322

/organism="Oryza sativa (indica cultivar-group)"

/mol_type="genomic DNA"

/db_xref="taxon:39946"

/clone_lib="Oryza sativa Expressed Library"

/notes="Oryza sativa exon trapped genomic sequences"

ORIGIN

Alignment Scores:

Pred. No.: 0.362 Length: 2322

Score: 115.00 Matches: 91

Percent Similarity: 35.71% Conservatives: 74

Best Local Similarity: 19.70% Mismatches: 187

Query Match: 2.97% Indels: 110

DB: 10 Gaps: 19

US-10-783-417-2 (1-735) x CL960568 (1-2322)

CL960568 2322 bp DNA linear GSS 21-SEP-2004
 OsIFCC004776 Oryza sativa Expressed Library Oryza sativa (indica
 cultivar-group) genomic, genomic survey sequence.

CL960568

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Chen Chen

Department of Bioinformatic

Beijing Institute of Genomics

Chinese Academy of Sciences, Beijing 101300, China

Tel: 86-10-80481559

Fax: 86-10-80488676

Email: chenchen@genomics.org.cn

Rice genomic sequence.

Class: exon-trapped.

Location/Qualifiers

1..2322

/organism="Oryza sativa (indica cultivar-group)"

/mol_type="genomic DNA"

/db_xref="taxon:39946"

/clone_lib="Oryza sativa Expressed Library"

/notes="Oryza sativa exon trapped genomic sequences"

ORIGIN

Alignment Scores:

Pred. No.: 0.362 Length: 2322

Score: 115.00 Matches: 91

Percent Similarity: 35.71% Conservatives: 74

Best Local Similarity: 19.70% Mismatches: 187

Query Match: 2.97% Indels: 110

DB: 10 Gaps: 19

US-10-783-417-2 (1-735) x CL960568 (1-2322)

Qy	256	IleGluProAsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIlePro	275
Db	628	TTAGAATTTGCAGCAGGACAAACAGAGAACTGGGCTGAGTTCCGGAACGCAAGTAGACAGT	687
Qy	276	LysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuAspGlu	295
Db	688	AGGATCTTAACTCGCTCAAGGATAGTTGCGAACAGTATTGGATGCAATTGCTGATGAC	747
Qy	296	ProAsnMetLysTrpSerIle---PheAsnAspTyrArgArgTyrMetThrIleThrVal	314
Db	748	AACTTGCAGAGATTACCGTGTTTCGAATATACAGAAATATTCTAAGATCATGCATCTCGTTGCA	807
Qy	315	LeuAspThrIleSerGlnPheSerLeuTyrAspIle-----	326
Db	808	CTTCTCACTGATGAAGATTGTTTTCAGAAAAATAGTCTTCTGTTGTGCCAAATA	867
Qy	326	-----	326
Db	868	ACTACTGAATCTTGTAATATATCAGTGCTCCCCCTACTAATGTGGGTGCTCATTCGAAT	927
Qy	327	LysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIle-----LysAsnGlu	344
Db	928	AAAAGTTTAGAAGTTTCATGGGCCCAAGTTGATTACCGAAATATATCATCTTTACAATGAG	987
Qy	345	LeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeuArgVal	364
Db	988	CTGGAGGAGCGTGTTAAAGAGTCATCATTTGGAAATGTGCACATGTGCGATCTCTCAAAA	1047
Qy	365	GlnProAsn-----LeuAlaThrMetGluTyr	373
Db	1048	TGTCCAATTCACATATTTAGCACTGCAATTCGACTGGTGTGAAGTTGGCATTTCCAGAAATG	1107
Qy	374	AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr	393
Db	1108	AGGATATTACAGAGCTTCCTTTGTTGTGACATTTTCAGTTTCGAAAGACTGTTTATATCTGCTA	1167
Qy	394	GluAsnThrAsnPheGlyAsnArgLeu---ValGlyIleSerAsnArgAspAlaProThr	412
Db	1168	CTGACTTGTCCATGGATTAATGAATTTGTTCTATTATACGACATTTATAGCACAA	1227
Qy	413	TyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySerProThrThrLys	432
Db	1228	ATGCATTGAGTTATCTCTCCAGATTGAAGGACGG-----GGCGCAGTGAAGCCAAAG	1281
Qy	433	ThrIleArgProPheGluSerTyrLysValSerIleValThrAspArgGlnSerProPro	452
Db	1282	CTCACAGATATTAT-----GCACAGGATCCGCTA	1311
Qy	453	ValSerProIleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySer	472
Db	1312	TGTGATACCAACAACTGAACTCATATATCTCAATATATCAAAATGATTTGGAGGCAGA	1371
Qy	473	SerAsnAsnThr-----LeuLysTyrSerAlaGlyGlySerLeuSerAsnTyr	498
Db	1372	AATGATATTACTGATGGGACCTGTTCAAGATATCCATTCTGAAGAACTCTCTTTTGTAT	1431
Qy	489	GlnAsnThrThrPhePheGlnPheProArgLysLysAspCysAsnLeuValIleAspPro	508
Db	1432	ATAACATT-----AAAAATTGCACCTATTTCACAGATGAT	1467
Qy	509	GlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThr	528
Db	1468	GGTATT-----GGTTATCGCGCTATGCTCTCAGCTAATGATGTAATGTAATA-----	1515
Qy	529	TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTrpThr	548
Db	1516	ACAAACTTCCTATGTTTAAAGGGAGACAACACTTACCGATCGCGCTCTCTCCAACTTTGTT	1575
Qy	549	HisSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIle-----	566
Db	1576	GGCTCTTCACCTGAG---TTCTTGATATTTCGAGACCGGTGGTTTCCATGGTATCATTTA	1632
Qy	567	---ProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGluGlyProGly	585

[illegible]

```
gene 1..2460
/ gene="At5g20250"

ORIGIN
Alignment Scores:
Pred. No.: 0.397 Length: 2460
Score: 115.00 Matches: 170
Percent Similarity: 31.70% Conservative: 109
Best Local Similarity: 19.32% Mismatches: 269
Query Match: 2.97% Indels: 332
DB: 49 Gaps: 49

US-10-783-417-2 (1-735) x CNS09YP2 (1-2460)

QY 57 ThrGlnTyrGlyAspAsnPheGlu-----ThrPheAlaSerAlaAspThrIle 72
Db 116 ACGAGGTACAGAGAAACGAAGAAGAGAGATATGACGATTAAACCGCG----- 166
QY 73 AlaAlaValSerAlaGlyThrIleVal-----SerGlyThrLeuLeuAlaGlyIleGly 90
Db 167 GTTCGTATTTCGACGGGAATCTAATCATCAAAACCGGACGATCTTAACCGGTGTACCA 226
QY 91 Gly-----LeuThrSerIleSer-----GlyProIle-----GlyIle 103
Db 227 GATAATGTCATCAGCAGCTGACGATCGGAAGCTGGACCGTAGAAGAGCTTCGTCGGA 286
QY 104 AlaIle-----IleIleSerPheGlyThrLeuIleThrVal 115
Db 287 GCTGTATTAAACAAGAGAGAGATTAACATCGTACCGATCGTACGCTTCGCAATTC 346
QY 116 PheTrpProAlaGlyGluGlnAspIleThrValTrpThrGlnPheIleIleMetGlyGlu 135
Db 347 CGGTATTATCTGTGTTCCGGTTTAAAG---CTCTGGTGGATGGCCAGAGAAATGGGAGAA 403
QY 136 IlePheValAspThrPro-----LeuThrGluSer----- 145
Db 404 ATGGCCGAGATATTCCTACGAGACGCAGATTCTTATTGTCGAGAGCAACGATGGGTCC 463
QY 146 -----IleLysGlnLeuLys 150
Db 464 CACCTTGAGTCTGACGGAGCTAACGGCGTCGAGTGAACCAAGAAAGTTTACACCGTTTC 523
QY 151 LeuGlnThrLeuGluGly---PheArgGlnIleLeuGln----- 162
Db 524 TTGCGGTAAATCGAGGATCTTTCGCTCGTCTCCAGAAACGTTAACGATGAGGTT 583
QY 163 -----SerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLys 175
Db 584 GAGCTCTGTTTGGAGAGTGGTGCAGCTGAGACACTAAACCGGTCGCTTACTCACTCTCTG 643
QY 176 ArgLeuGlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeu 195
Db 644 TATATTACGCC---GGTACAGATCCGTTTCAGACAATTAACGGACGCTATTTCGACCGGT 700
QY 196 LysIleArgPheGluAsnValHisAsnAspPheIleArgGluIleProGly----- 212
Db 701 AGTTGCATCTGAAATAGTTTCGCTCAACGTCACGAAAGAAAGAGCTTCAGGGATCGTTGAC 760
QY 213 -----PheGlnLeuGluThrTyrTrpLysThrLeuLeuLeuProIleTyrAlaGlnAlaAsn 231
Db 761 TACTTGGATGTGTGCACTTGGGACGCGTTTATATCAAGAAGTGAATCAGGAAGCGTCGAA 820
QY 232 PheHisLeuAsnLeuGlnGlnGlyAlaGlu-----LeuAlaAsp 245
Db 821 GCTGGTCTTAAGTCTCTCGCGCGCGGTGTGACGCCCGGAAGTTTGTATTATACAGACGAC 880
QY 246 GluTrpAsnAlaAspIleHisProSerGlnIleGlu----- 257
Db 881 GGTGGCAATCAGTTGAGAGAGATGCCACGGTGGAGCCCGGAGATCAGAGAAGAGATCA 940
QY 258 -----Pro 258
Db 941 CCGATTTTTCGACTGACGGGGATCAAGGAGAACGAAAGTTTAAAGAAGGATGATCCA 1000
```

```
QY 259 AsnAlaGlyThrSerAspAspTyrTyrTyrLysLeuLeuLysGlu----- 272
Db 1001 AACGTTGGGATTAAAG---AACATTTGTCAAGATTCTCAAGGAGAAACACGGTTTGAATAT 1057
QY 272 ----- 272
Db 1058 GTTTACGTGGCGACGCGATACCGGTTATTGGGCGGAGTTAGACCGGGGAGAAATAC 1117
QY 273 ---AsnIleProLysTyrSerAsnTyrCysAlaAsn-----ThrTyr 285
Db 1118 GGATCTGTTATGAATATCTTAATATGTCAAAGCGGTGTGGGAGATGATCCACGCTGG 1177
QY 286 ArgThrGlyLeuLysAsnLeuArgAsp-----GluProAsnMetLysTrpSer 301
Db 1178 AAGACTGATGATTAATGACGCTGCAAGGTTGGTGGTTAGCCCAAGAAAGTGTATAAG 1237
QY 302 IlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPhe 321
Db 1238 TTTTACAATGAGCTTCATAGTTACTTGGCTGATGCTGGCGTGGACGGTGTGAAA----- 1291
QY 322 SerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIle 341
Db 1292 -----GTGGATGTGAGTGTGATTGGACATTGGGTGGT-----GGTTTAGCGCGT 1339
QY 342 LysAsnGluLeuThrArgGluIleTyr-----ThrThrGluIleAsnPhe 356
Db 1340 CGGGTTGAGCTGACTCGTCAAGTTTCAAGCTCTTGACTCTCTGTTGCTAAGAACTTC 1399
QY 357 AspArgLeuProGlnLeuArgValGlnProAsn-----LeuAlaThrMetGluTyr 373
Db 1400 -----CCGATAAATGGCTGCATTTGCTGTATGATGAGCCAT 1432
QY 374 Asn-----LeuThrArgAlaSerPheLysLeu 382
Db 1433 AATACGGAGTCTCTTACTGCTCGAAGCAAGAGCTGTGATTAGATGATGATGATGAT--- 1489
QY 383 PheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsnArgLeu 402
Db 1490 -----TTCATATCCA----- 1498
QY 403 ValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyr 422
Db 1499 -----CGGATCGGCTGCTCATACCATCATATAGCTCGCTGCTTAT 1543
QY 423 -----GlyGlu-----ArgThrGlySerProThrThrLysThrIleArg 435
Db 1544 AACAGTGTGTTCTGGGAGAGTTTATGACGCTGACTGGGATATGTTCCATTCCTGTCAC 1603
QY 436 ProPheGluSerTyrLysValSerIleValThrAspArgGlnSerProProValSerPro 455
Db 1604 CTTGCTGCAGAGTATCATGCTCTGCTAGGGCCATCATGCTGGTGGCT----- 1651
QY 456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn 475
Db 1652 -----CTCTATGTCAGTGTATCTCCTGGA--- 1675
QY 476 ThrLeuLysTyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGln 495
Db 1676 -----AAGCACAAC-----TTTGAG 1690
QY 496 PheProArgLys----- 499
Db 1691 CTTCTAAGAAAGTGTATTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1750
QY 500 -----LysAspCysAsnLeuValIleAspPro-----GlyCysSerPro----- 512
Db 1751 CCAACTCGCGATGTT---TTGTTTCGCTGATCTCTCCCGTGCATGCTGCTGCTGCTGCTGCT 1807
QY 513 -----AsnPheAsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSer 530
Db 1808 ATATGGAACATGAACAGTAC----- 1828
```

```

QY 531 TyrValIleGlyLeuGlnLeuGlnLeuAspThrGlyValLeuGly----- 546
Db 1829 -----ACTGGAGTCTTGGCGGTATATACTGC 1855
QY 547 -----TrpThrHisSerSerValAspArgTyrAsnAlaIle----SerAspLys 561
Db 1856 CAAGGAGCAGCTTGG-----AGCAGCACAGAGAGAAAAAATTTTCCACGACTAAA 1909
QY 562 IleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIle 581
Db 1910 ACTGATAGCTCTACTGCTCATTCGGTGGTGTGATGCTGATTCATTAATATCGAGGCCCTCC 1969
QY 582 GluGlyProGlyHisThrGlyGlyAsn---LeuValTyrLeuGlnSerGlnGlyArgLeu 600
Db 1970 ACTGATCCACACACCTGATGAGACTGCTGCTTTACTCCACAGACAGAGCGCACTT 2029
QY 601 GluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAla 620
Db 2030 ATTGTTATGCCATACATATGCTCTCTCCAGTCTCACTCAAAATCCGTGAGCAGCATC 2089
QY 621 ThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGly 640
Db 2090 TTC-----ACGGTGAGGCCCATTAGTCATCTGTTGATGTTGATCTCTTTT 2134
QY 641 IleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyr 660
Db 2135 GCCCAATTGCTCTAGTAAACATGTACAAATTCGGAGGAGCTATCGAAGACTTAGATAT 2194
QY 661 -----GlyAspPheGlyTyrPhe 666
Db 2195 GAAGCCGAGAGATGAAGTGGTAATGAAGTTAAAGATGTGGCAAAATTCGGATCTTAC 2254
QY 667 GlnPhe-----ProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePhe 684
Db 2255 TCTTCTGTGAAGCTAAGAGATGGTGTGTGAGTCAAAAT---GAGATTGCATTCGAGTAC 2311
QY 685 AsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIleGluPheIleProIle 704
Db 2312 GAT-----TCCTCTCTGGATTGGTCACCTTTGATAGACAAATAGCCCTATA 2359
QY 705 ThrSerSerMetHisGlnAsnArgGlyLysLeuGlnLeuThrIleGlnThrLysIle 724
Db 2360 -----GAGAACAAACGATTTTCATCTGATCCAAAGTTGAGTTA 2395

CF951415 820 bp mRNA linear EST 20-NOV-2003
LOCUS UI-M-HLO-cnb-o-16-0-UI_r1 NIH_BMAP_HLO Mus musculus cDNA clone
DEFINITION IMAGE:30633855 5', mRNA sequence.
ACCESSION CF951415
VERSION CF951415.1 GI:38467284
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers

```

FEATURES

```

source
1..820
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone IMAGE:30633855"
/tissue_type="Upper Head"
/dev_stage="9.5 dpc"
/lab_host="DH10B (TI phage resistant)"
/clone_lib="NIH_BMAP_HLO"
/note="Organ: Head; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CGAATCGAAT. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH)."

ORIGIN
Alignment Scores:
Pred. No.: 0.0764 Length: 820
Score: 114.50 Matches: 54
Percent Similarity: 44.26% Conservative: 50
Best Local Similarity: 22.98% Mismatches: 106
Query Match: 2.96% Indels: 25
DB: 7 Gaps: 10

US-10-783-417-2 (1-735) x CF951415 (1-820)
QY 188 LeuGlnGlnAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPheIle 207
Db 51 CTCACCTCTCCCTCGCTGATGTCGCCGACGATATGAAGCGTGGTCCCAAGAACCTC 110
QY 208 ArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeuProIleTyrAla 227
Db 111 CAGAGGCCGAG-----GAATGTACAAAGTCCAAGTTGCTGACCTC---TCT 155
QY 228 GlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTrp 247
Db 156 GAGGCTGCCAACCGAACAACGATGCCCTGGCCGAGCCAGCAGAGACTCAACAGTAC 215
QY 248 AsnAlaAspIleHisProSerGlnIleGluProAsnAla-----GlyThrSerAsp 265
Db 216 CGGAGACAGGTGCAGTCACTCACCTGTGAAGTGGATGCCCTTAAAGGCACCTAACGAGTCC 275
QY 266 TyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyr 285
Db 276 CTGAGCGCCAGATGCGTGGATGGAAGAAATTTGCCCTTGAAGCTGCTAAC---TAC 332
QY 286 ArgThrGlyLeuLysAsnLeuArgAspGlu---ProAsnMetLysTrpSerIlePheAsn 304
Db 333 CAGGACACTATTGGCCGCTGCAGATGAGTCCAAACATGAGGAGAGAGATGCTCGT 392
QY 305 AspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyr 324
Db 393 CACCTTCGTGAATACCAAGATCTCTCAATGTTAAGATGGAAGAGAAATTTTGGCCTT--- 449
QY 325 AspIleLysArgTyrArgAspSerIleGlyIle-----GluValLysGlyIleLys 342
Db 450 GAAAGCTCTAACTACCAAGACACTATTGGCGCTCGCAGGATGAGATCCANAACATGAAG 509
QY 343 AsnGluThrThrArgGluIle-----TyrThrThrGluIleAsnPheAspArgLeuPro 360
Db 510 GAAGAGATGCTCTGCTACCTTCGTCGTAATCAAGATCTGCTCAATGTTAAGATGCCCTG 569
QY 361 GlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPhe 380

```



```

Db      226 ACCCCCTTTTGGGAAAAAACAACCTCTCCCTTCACTA-AACCTACTTCTTT 168
Qy      681 -----proPhelePheAsnArgAlaAspValSerAsnSerIleLeuI 695
Db      167 AATAAAACCCGCCACCTTTTGG-----AAAAATACCTTCTCTCC 123
Qy      695 leileAspLysIleGluPheIleProIleThrSerSerMet 708
Db      122 GCCCTCGCCCTTTTATTATACACCCATTACTCTCTCCCTC 82

RESULT 14
CL974457      2841 bp DNA linear GSS 21-SEP-2004
LOCUS      OsIFCC025896 Oryza sativa Express Library Oryza sativa (indica
DEFINITION      cultivar-group) genomic, genomic survey sequence.
ACCESSION      CL974457
VERSION      CL974457.1 GI:52403441
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE      1 (bases 1 to 2841)
AUTHORS      Ma,L., Wang,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
Wong,K.S., Deng,X.W. and Wang,J.
TITLE      An analysis of transcriptional regulation of the rice genome and
JOURNAL      its comparison to Arabidopsis
COMMENT      Unpublished (2004)
Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
FEATURES
source
1..2841
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/notes="Oryza sativa exon trapped genomic sequences"

ORIGIN
Alignment Scores:
Pred. No.:      0.982      Length:      2841
Score:      112.50      Matches:      175
Percent Similarity:      33.06%      Conservative:      111
Best Local Similarity:      20.23%      Mismatches:      313
Query Match:      2.91%      Indels:      267
DB:      10      Gaps:      47

US-10-783-417-2 (1-735) x CL974457 (1-2841)
Qy      24 ArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGln-----ProLeu 41
Db      499 CGCAATTGAAGACGTTGCCAAGAAAATTTGCACACATATAAGCTTAATCCCTG 558
Qy      42 GlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGlyAsp 61
Db      559 GACTCTACC-----ATGTTGGGAATATATAAGCAACGCGAGT 597
Qy      62 AsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleVal 81
Db      598 AATTAGACGTAGAGGTGTAAGACCAATTCGAGCAAGTC-----GGATGGGTCAAT 651
Qy      82 SerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIle 101
Db      652 GCC-----AAGAAATGTGAGGCGATAGCATTTAGCGCCCATGCACTTGGTTTC 699

```

```

Qy      102 Ile-----GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAla 119
Db      700 TTGCTATCTGAATGGACCTTTGTGAATGGAGACACTG----- 738
Qy      120 GlyGluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAsp 139
Db      739 -----AGCAATAGTAGATATTTGGACT-----GAAGCTTTTCTGTAT 774
Qy      140 ThrProLeuThrGluSerIleLysGlnLysLeuGlnThrLeuGluGlyPheArgGln 159
Db      775 AATTCAGCTGTGGCCATCTCTGAAA---CTAACTTACAAGAATATGCCCACTATTATTGAGA 831
Qy      160 IleLeuGlnSerTyr-----AsnThrAla----- 167
Db      832 TTATGCTTTTGCCTACTCGCCCATCTTCCCTAAAGGTCAATATAGCTAAAGCAAGTCTG 891
Qy      168 LeuAspAspTyrArgLysLeuLysArgLeuGlnAlaProGlyLeuProProSerSerAla 187
Db      892 GTTCACCAATGGATGCTCTTGGTTTATTGAG-----CCATCTAAAACA 936
Qy      188 LeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPheIle 207
Db      937 TTTTCATCCGTACGACTTGGCGAGAG-----TATATC 969
Qy      208 ArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeuProIleTyrAla 227
Db      970 AGACACACTTGTGGG-----ATGCTTTCTCCTCAGCGTTCATATTCACACTGAACAA 1023
Qy      228 GlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTrp 247
Db      1024 GAACTTTTCCACCATGCACGATATGTTTCATGATGTCGAGATCGGTATGATGAAGAG 1083
Qy      248 AsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTyrTyr 267
Db      1084 TTGGTTTTCTTAATGACACAAAATTTAGCAGTACCAGACAGAACAAAATTTTGCACATAT 1143
Qy      268 LysLeuLysGluAsnIleProLysTyrSerAsnTyr----- 280
Db      1144 GCTTTGCTT---GAAAATTATAGCAAGTCATCAATCTATCAACTATATATTCGTCGCCAG 1200
Qy      281 -----CysAlaAsnThrTyrArgThrGly----- 288
Db      1201 TTAAGGCGCAGTACATACTTCGAATTGCAGCAAAATTTGGTGTCCAAAGGTGATGAATTTCA 1260
Qy      288 ----- 288
Db      1261 TTCACAAAGTTCTTGGTGTCTTAGATTAACTGCTGCTCCATAAGGATCCTCCCAAGC 1320
Qy      289 -----LeuLysAsnLeuArg-----AspGluProAsnMetLysTrpSerIle 302
Db      1321 TCTATTGGCAAAATGAAGCAGCTGAGGTTTCTTATCGCTCCCAACATAGGAGATAATGTG 1380
Qy      303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db      1381 TTCCCA-----AAGAGTATTACTCTGCTCCCAAACTA----- 1413
Qy      323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyIleGluValLysGlyIleLys 342
Db      1414 -----AAATACCTTGGCTCCATCGATCTTTTCGAATTTTCAGCATTTGCAA 1458
Qy      343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362
Db      1459 GGATCAATTAGCAACATGCTTGTCTGNTACATCTTGACCTATCTGGTGTCTCTAATATC 1518
Qy      363 ArgVal---GlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPheLys 381
Db      1519 AGATTATACAGCCAGAGAGCTCTATGT-----GSCCTCACCAAACTGCAGTTTGTG 1569
Qy      382 LeuPheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsnArg 401
Db      1570 AACTTATCGTGGTCTCCATCTTCTCAAAATACTGCAGAAAATATTGCAAGTCTTACTGAA 1629

```



```
QY 402 Leu-----ValGlyIleSerAsn-----ArgAspAlaProThrTyrSerAsn 415
|||
Db 1630 CTGCAATATTAACTTAACTCAAACTCCTCCCTCTTCAGCCAGGTATACCATCTCATATTGGG 1689
416 ThrIleThrGluThrLeuTyrGlyGluArgThrGly---SerProThrThrIleThrIle 434
:
Db 1690 AGCCTCACTGAACACTAAATATTGAATTTAGCTTATCAGGTGTTCAGGGCCCTTGTAATAATTGCCA 1749
435 ArgProPheGluSerTyrIlysValSerIle----- 444
|||
Db 1750 ATGTGCTTTCGGAATCTGAAGAATTGGTGATACATCTTGACCTGTGCAGGCTGCTCCGGGGTA 1809
445 -----ValThrAspArgGlnSerProProValSerPro 455
:
Db 1810 CAAGATTTCAAAACAAGTCTTTGGTGGGCTCACAAACTGCAGTACTTAAACTTATCAAAG 1869
456 Ile-----GlnProHisPheIle----- 461
|||
Db 1870 ATATTGGTAGAACACAGAGTGGTGATTAATTGGGATGGTTATCTCGAACTATCAGTACT 1929
462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsnThrLeuIlysTyr----- 479
:
Db 1930 CTTAATGATCTTGAG---TATCTAAACCTTCTCGGAATAGCAGGATTGATTATCTACCA 1986
480 SerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArgLys 499
:
Db 1987 AGAAGCCTTGGAAACCTCAAAAAGCTTCAA---ACATTGGACCTTTTCATACCTGCGGTCT 2043
500 -----LysAspCysAsnLeuValIleAspPro 508
2044 CTGAGATCCTCCACACAGCATAGATAATTAATGACAGTCTGGAGTCTCTAATAGTTGTG 2103
509 GlyCysSerProAsnPheAsnAsnTyr-----Ser 518
2104 GGGTCTCAGATCAACTTAAGAGTATTTGAGAAATCTCAATTCAGAATATTCAGTCT 2163
519 HisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIle-----Gly 534
2164 CATTAC-----CACACTTCATTTGTTCACTCTATTTCTGTCCATAGAAGAAGCAGGAG 2217
535 LeuGlnLeu-----GlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSer 551
2218 AVTGAGTTGTCAGAGAAACAAATCTTTCAACATTGCACATTTTCATTGGACT---TCAAGG 2274
552 ValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIlePro----- 567
2275 GCTGACCGACTTTGGAGACAGAGATGTTTGGGAGAACTAATGCCACCAAGAGTCTG 2334
568 -----AlaIleLysGlyAsnLeuAspThrAsnSerLysValIleGluGlyPro 584
2335 TGGAAATTTAAGCATACAAGGGTATGATAGACACACCTTCCCACTTGGTTGTGAGCATC 2394
585 GlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIleThrCys 604
2395 TCACATCATCTTCCCAAGTCTTGTAGATTGAGTT--GAAGGATTTCACAGATG----- 2447
605 GluThrProAsnSerThr----- 610
2448 CAGACAGTCTGCCACCACTGGGACAGTTCGCAACCTGAATGAAGTGATCTCCACAGAT 2507
611 ---GlnSerTyr-----PheIleArgLeuArgTyrAlaThrAsnGlyAla 624
2508 GCAGACTCTCACAAAATTTGATGGCGGTTTCTCGCGTGGGACAA----- 2552
625 GlyAsnThrLeuProAsnIle---SerLeuThrIleProGlyValIleGlyIleProPro 643
2553 ---AGCGCATTCGCAAACTGAAAGTCTTCACTTTCAGGGATAT----- 2594
644 GlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGln-TyrGly----- 661
2595 -----GAACAGTTGGAGAGTGGAGCACCACACACAGTAAAGTGTGTA 2642
662 -AspPheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsn----- 677
```

```
Db 2643 GGATAGCATAGAGTTCATGTTCCCTATGCTAGTTTACACTGTCAATAGAGCTTTTCCAAAT 2702
QY 678 -----ArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIleLe 694
:
Db 2703 GCTTAGACTGAACACCATGCCACCGATGTTTCACAGCTGGCTCATATCAGAGAGCGCAA 2762
694 uilleIleAsp-----LysIle-----GluPheIlePr 703
2763 GGCTGTGCGACCGGTTTAACGGAGGCTCTGATGAAGATTTCGCTGAAGATTTCATGCC 2822
QY 703 olleThrSerSer 707
Db 2823 TACCGACTCTTCA 2835

RESULT 15
CNSOALJE 1143 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLTFB30ZG08 of Flowers and buds of strain col-0 of Arabidopsis
thaliana (thale cress).
ACCESSION BX829716.1 GI:42458466
VERSION HTC; GSUT cDNA.
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 1143)
AUTHORS Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1143)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
Genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES
Location/Qualifiers
1..1143
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLTFB30ZG08"
/tissue_type="Flowers and buds"
/scotype="Col-0"
/plasmid="pCMVSPORT_6"
complement(1..1143)
/gene="At5g56640"

ORIGIN
Alignment Scores: 0.335 Length: 1143
Pred. No.: 58 Matches: 58
Score: 111.00
Percent Similarity: 37.33%
```

Best Local Similarity: 26.73% Mismatches: 86
Query Match: 2.87% Indels: 50
DB: 4 Gaps: 11

US-10-783-417-2 (1-735) x CNS0A1JE (1-1143)

```
QY 139 AspThrProLeuThrGluSerIleLeuLeuLysGlnLeuLysLeuGlnThrLeuGluGlyPheArg 158
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 380 GATCCCGACCTCGACGAAACACAAATTCACATTTGGTTTCAATCTGCGGAAGCAATCGGT 439
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 159 GlnIleLeuGlnSerTyrAsn-----ThrAlaLeu---AspAspTrpArgLys 173
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 440 AAAGATTATCTTAACGAAATTTGGCTTCATCTACCTGCGCTTATCCATGATCTTTGGGAAG 499
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 174 LeuLysArgLeuGlnAlaProGlyLeuProSerSerAlaLeuGlnGlnAlaAlaLeu 193
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 500 GTTCTTACCCTACACAGTTTGGAGGCTTCCTCAATGGGCTGTTTGGTGACACATTT 559
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 194 ThrLeuLysIleArgPheGlu-----AsnValHisAsnAspPheIleArgGluIlePro 211
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 560 CCTGTGGGTGTCATTTGATGATCTAAGTACACACAGTATTTTATGGAACCCCT 619
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 212 GlyPheGlnLeuGluThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsn 231
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 620 GATTTTAAACAACCTTAAGTACATACC--AAAGCTGGAATTTATTCTGAAGGCTGTGGA 676
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 232 PheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIle 251
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 677 CTTGAAAAATGTTCTTATGTCATGGGA----- 703
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 252 HisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTyr---TyrLysLeuLeu 270
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 704 -----CATGATGACTACATGATCTCGTGGCC 730
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 271 LysGluAsn-----IleProLys----- 276
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 731 AAGGAGATGGAAGCACCTTACCATCCCGGGACTATTATCATCCGATACCATTCCTTT 790
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 277 TyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro 296
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 791 TATCCATTGTACAAGGCTGGAGCTTAC--ACTCACCTTATGAACGAGGAAGACAAGGAG 847
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 297 AsnMetLysTrp---SerIlePheAsnAspTyrArgTyrMetThrIleThrVal--- 314
   ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : |||
Db 848 AATCTCAAAATGGCTTCATGTTTTTAAACAAGTATTACTTATACAGCAAGACCAAGTACAC 907
   ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : |||
QY 315 -----LeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyr 329
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 908 GTTAATGTCGAGAAGGTTAAGCCGTACTACATTTCTTATCAAGAAATAT 958
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: December 12, 2005, 04:21:08
Job time : 7319 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 11, 2005, 14:08:26 ; Search time 332 Seconds
(without alignments)
3935.265 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869

Sequence: 1 MNQNDNNEVEIIDSHTSPY.....KLETIQKINTFFNHTKTL 735

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10783417/runat_01122005_141954_24196/app_query.fasta_1.903
-DB=Issued Patents NA -OFMT=fastaf -SUFPFX=tni -MINMATCH=0.1 -LOOPECL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=spt -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10783417 @CNG 1 1 200 @runat_01122005_141954_24196 -NCPV=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1060.5	27.4	3543	3	US-09-224-024-27
2	1060.5	27.4	3543	6	PCT-US94-07902-27
3	899.5	23.2	2061	3	US-09-224-024-30
4	899.5	23.2	2061	6	PCT-US94-07902-30
5	775.5	20.0	3507	2	US-08-315-468-3
6	676	17.5	1953	2	US-08-315-468-5
7	676	17.5	1953	3	US-07-941-650A-3
8	663.5	17.1	2430	6	PCT-US92-00040-1
9	663	17.1	1959	3	US-08-993-775B-37

10	663	17.1	1959	3	US-08-993-722A-37	Sequence 37, Appl
11	663	17.1	1959	3	US-08-993-170A-37	Sequence 37, Appl
12	663	17.1	1959	3	US-08-993-775B-37	Sequence 37, Appl
13	663	17.1	1959	3	US-09-427-770-37	Sequence 37, Appl
14	663	17.1	1959	3	US-09-427-769-37	Sequence 37, Appl
15	660.5	17.1	3414	2	US-07-973-320-3	Sequence 3, Appl
16	660	17.1	1959	3	US-09-377-466B-3	Sequence 3, Appl
17	660	17.1	1959	3	US-10-232-665-3	Sequence 3, Appl
18	658	17.0	1959	3	US-08-996-441B-27	Sequence 27, Appl
19	658	17.0	1959	3	US-08-996-441B-35	Sequence 35, Appl
20	658	17.0	1959	3	US-08-993-722A-27	Sequence 27, Appl
21	658	17.0	1959	3	US-08-993-722A-35	Sequence 35, Appl
22	658	17.0	1959	3	US-08-993-170A-27	Sequence 27, Appl
23	658	17.0	1959	3	US-08-993-170A-35	Sequence 35, Appl
24	658	17.0	1959	3	US-08-993-775B-27	Sequence 27, Appl
25	658	17.0	1959	3	US-08-993-775B-35	Sequence 35, Appl
26	658	17.0	1959	3	US-09-427-770-27	Sequence 27, Appl
27	658	17.0	1959	3	US-09-427-770-35	Sequence 35, Appl
28	658	17.0	1959	3	US-09-427-769-27	Sequence 27, Appl
29	658	17.0	1959	3	US-09-427-769-35	Sequence 35, Appl
30	656	17.0	1962	3	US-09-377-466B-5	Sequence 5, Appl
31	656	17.0	1962	3	US-10-232-665-5	Sequence 5, Appl
32	656	17.0	2050	3	US-08-996-441B-101	Sequence 101, App
33	656	17.0	2050	3	US-08-993-722A-101	Sequence 101, App
34	656	17.0	2050	3	US-08-993-170A-101	Sequence 101, App
35	656	17.0	2050	3	US-08-993-775B-101	Sequence 101, App
36	656	17.0	2050	3	US-09-427-770-101	Sequence 101, App
37	656	17.0	2050	3	US-09-427-769-101	Sequence 101, App
38	655	16.9	1959	3	US-08-996-441B-17	Sequence 17, Appl
39	655	16.9	1959	3	US-08-996-441B-59	Sequence 59, Appl
40	655	16.9	1959	3	US-08-993-722A-17	Sequence 17, Appl
41	655	16.9	1959	3	US-08-993-722A-59	Sequence 59, Appl
42	655	16.9	1959	3	US-08-993-170A-17	Sequence 17, Appl
43	655	16.9	1959	3	US-08-993-170A-59	Sequence 59, Appl
44	655	16.9	1959	3	US-08-993-775B-17	Sequence 17, Appl
45	655	16.9	1959	3	US-08-993-775B-59	Sequence 59, Appl

ALIGNMENTS

RESULT 1

US-09-224-024-27
; Sequence 27, Application US/09224024
; Patent No. 6056953
; GENERAL INFORMATION:
; APPLICANT: Leslie Hickie
; APPLICANT: Jewel Payne
; TITLE OF INVENTION: Materials and Methods for the Control of
; TITLE OF INVENTION: Calliphoridae Pests
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/224,024
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/856,226
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA79

TELECOMMUNICATION INFORMATION:

TELEPHONE: 904-375-8100

TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 3543 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-09-224-024-27

Alignment Scores:

Pred. No.: 2,54e-102 Length: 3543
 Score: 1060.50 Matches: 271
 Percent Similarity: 50.26% Conservatives: 117
 Best Local Similarity: 35.10% Mismatches: 279
 Query Match: 27.41% Indels: 105
 DB: 3 Gaps: 24

US-10-783-417-2 (1-735) x US-09-224-024-27 (1-3543)

QY 1 MetAsnGlnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
 DB 1 ATGAATCCTTATCAAAATAAATAATGATATGAAACATTAAATGCTTCACAAAAAATTA 60
 QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
 DB 61 -----AATATATCTAATAATATATACAAAGATATCCAAATAGAAATAGTCCAAACAAATTA 114
 QY 41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
 DB 115 TTACAAAGTACAAATATATAAGATTTGGCTCAATATGTGTCAACAGATCAGCAGTATGGT 174
 QY 61 AspAsnGluThrPheAlaSerAlaAspThrIleAlaIleValSerAlaGlyThrIle 80
 DB 175 GGAGATTTTGAACCTTTTATTGTAGT-----GGTGAACCTCAGTGCCTATACTATT 225
 QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
 DB 226 GTAGTTGGGACCTACTGCTGTTTCGGGTTCACAACA-----CCCTTAGGA 273
 QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGly 120
 DB 274 CTT-----GCTTTAATAGTTTGTGTACATTATACAGTCTCTTTCCAGCCCAA 324
 QY 121 GluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
 DB 325 GACCAATCTAACACA---TGGAGTGTCTTTATATACACAAACTAAATAATATTATATAAAAAA 381
 QY 141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
 DB 382 GAAATAGCATCAACATATATAGTAATGCTTAATAAATAATTTTAAACAGGTCGTTAATGTT 441
 QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaPro 180
 DB 442 ATCAGCAGCTTATCATATATACCTTTAAACATGGAG-----AATAATCCA 486
 QY 181 GlyLeuProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGlu 200
 DB 487 AACCCCAAAATACTCAGGATGTAAGGACACAAATCCAGCTAGTTTCATTACCATTTTCAA 546
 QY 201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr 218
 DB 547 AATGTCATTCAGAGCTTGTAAACTCTTGTGCTCTTAATCCTAGTATGCGATTACTAT 606
 QY 219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAsnPheHisLeuAsnLeuGln 238
 DB 607 AACATCATAGTATTATCTAGTTATGACAAAGCAGCAAACTTACATCTGACTGCTGATTAAAT 666
 QY 239 GlnGlyAlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluPro 258
 DB 667 CAAGCCGTCAAATTTGAGGCTATTTTAAAAAACAATCGACAATTCGATTATTATAGACCT 726

QY 259 AsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSer 278
 DB 727 ---TTGCCAACAGCAATTTGATTATATCCAGTATGCTAAAGCTATAGAAGATTACACT 783
 QY 279 AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro----- 296
 DB 784 AATTATTGTGTAAACAACCTTATAAAAAAGGATTAATTAATTAATAAACACGCGCTGATAGT 843
 QY 297 -----AsnMetLysTyrSerIlePheAsnAspTyrArgArgTyrMetThrIle 312
 DB 844 AATCTTGATGGAATATAAACTGGAAACACATACATACGTATCGAACAAAATGACTACT 903
 QY 313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer 332
 DB 904 GCTGTATTAGATCTGTGTGACCTCTTCTCTAATTAATGATGAGTAAATAT----- 954
 QY 333 IleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThr 352
 DB 955 -----CCAATAGGTGTCCAATCTGAACCTTACTCGAGAAATTTATCAGGTA 999
 QY 353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
 DB 1000 ---CTTAACCTTCGAAGAAAGCCCTTATAATATTAT-----GACTTCAATATCAAGAG 1050
 QY 373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyr 392
 DB 1051 GATTCACCTACAGCTAGACCG---CAATTATTACTTGGCTTGATCTTCTGAAATTTTAT 1107
 QY 393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
 DB 1108 GMAAAGCGCAACTACTCTCTAATAATTTTTCACGACCATTAATAATGTTTCATTAC 1167
 QY 402 ---LeuValGlyIleSerAsnArgAspAlaProTyrSerAsnThrIleThrGluThr 420
 DB 1168 ACATTGATAATATATCCCAAAATCTAGTGTGTTTGGAAATCAATGTAAGTAAATA 1227
 QY 421 LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440
 DB 1228 TTA-----AAATCTCTGTTGGCAACAATATTATATT----- 1263
 QY 441 LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
 DB 1264 -----TTT 1266
 QY 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
 DB 1267 TTATTAAATGTCATAAGCTTAGATAATAATAATATCTAAATGATTATAATAATTATTAGTAAA 1326
 QY 477 -----LeuLysTyrSerAlaGlyGly 483
 DB 1327 ATGATTTTTTTTATACTAATGCTAGACTTTTGGAGAAAGAACTTACAGCAGGATCT 1386
 QY 484 SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysLysAspCys 502
 DB 1387 GGGCAATAACTTATGATGTAATAATAATAATATTTTCGGTTACCAATCTTAAACCAAGA 1446
 QY 503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
 DB 1447 GAGAATCAAGCAATCCCTACCTTTTCCAAACATATGATACTATAGTCATATTTATCA 1506
 QY 523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThr 542
 DB 1507 TTTTATTAAGTCTTAGTATCCCTCGCAACATATAAAACTCAAGTGTAT----- 1554
 QY 543 GlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
 DB 1555 ---ACGTTTCTGGACACACTCTAGTGTGTGATCTTAAATAATACATTTATACACATTTA 1611
 QY 563 IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582
 DB 1612 ACTACCCCAATTCAGCTGTAAAGCGAATTCACCTTGGGACTGCTTCTTAAGTTGTTCAA 1671

Db 784 AATTATGTTGTAACAACTTATAAAAGGATTAATTAATAAAGCAGCGCTGATGT 843
QY 297 -----AsnMetIysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIle 312
Db 844 AATCTTGATGGAATAATAAAGTGAACACATACATACGTATCGAACAAAAATGACTACT 903
QY 313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleYsArgTyrArgAspSer 332
Db 904 GCTGTATTAGATCTTGTCACCTCTTCTCAATATTATGATGATTAATAT----- 954
QY 333 IleGlyGlyIleGluValIysGlyIleLysAsnGluLeuThrArgGluIleTyrThr 352
Db 955 -----CCAATAGTGTCCTCAATCTGAACTTACTCGAGAAATTTATCAGGTA 999
QY 353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
Db 1000 ---CTTAACCTCGAAGAACCCCTATAATATTAT-----GACTTTCATATCAAGAG 1050
QY 373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyr 392
Db 1051 GATTTCATTTACAGGTAGACCG---CATTTATTACTTGGCTTGATCTTTTGAATTTTAT 1107
QY 393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
Db 1108 GAAAAAGCGCAAACTACTCTTAATAATTTTCCAGCGCAATTAATAATGTTTCATTAC 1167
QY 402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
Db 1168 ACATCTGATAATATATCCAAAATCTAGTGTGTTTGGAAATCAATGTAATGATAAA 1227
QY 421 LeuTyrGlyGluArgThrGlySerProThrTyrLysThrIleArgProPheGluSerTyr 440
Db 1228 TTA-----AAATCTCTGTTGGTGGCAACAAATATTATATT----- 1263
QY 441 LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
Db 1264 -----TTT 1266
QY 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
Db 1267 TTATTAAATGTCATAAGCTTAGATAATAATAATATCTAAAGATTAATAATATAGTAAA 1326
QY 477 -----LeuLysTyrSerAlaGlyGly 483
Db 1327 ATGGATTTTTTATTAACATAATGCTAGTACTTTTGGAGAAAGAACTTACAGCAGATCT 1386
QY 484 SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysLysAspCys 502
Db 1387 GGGCAAAATAACTTATGATGTAATAATAAAATATTTTCGGGTGTACCAATTTCTTAAACCAAGA 1446
QY 503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
Db 1447 GAGAAATCAAGCAATCCCTACCTTTTCCAAACATATGATAACTATAGTCATATTTTATCA 1506
QY 523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThr 542
Db 1507 TTTATTAAGCTTAGTATCCCTGCAACATATAAACTCAAGTGTAT----- 1554
QY 543 GlyValLeuGlyTrpThrHisSerSerValAspArgTyrAsnAlaIleSerAspIle 562
Db 1555 ---ACGTTTGCTGGACACACTAGTGTGTGCTTAAATAACAAATTTATATACATTTA 1611
QY 563 IleThrMetIleProIleIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582
Db 1612 ACTACCAAAATCCAGCTGTAAAGCGAATTCATTTGGGCTCTCTTAAAGTTGTTCAA 1671
QY 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle 602
Db 1672 GGACCTGGTCATACAGAGGGGATTAAT-----GATTTCAAAGATCATTTCAAAT 1725
QY 603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622

Db 1726 ACATGTCAACACTCAAAATTTTCACAATCGTATTATTAAGAATTCGTATGCTTCAAT 1785
QY 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db 1786 GGAAGCGCAAAATACACGAGCTGTTTATAAATCTTAGTATCCAGGGGTAGCAGAACTG--- 1842
QY 643 ProGlnArgIleAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
Db 1843 GGTATGCACTCAACCCCACTTTTCTGTCAGATATACGAATTTAAATAATAAAGAT 1902
QY 663 PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
Db 1903 TTTCAAGTACTTGAATTTTCTAAGAGGTGAATTTGCTCCAAATCAAAACATATCTCTT 1962
QY 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleAspLysIleGluPhe 701
Db 1963 GTGTTTAAATCGTTCGGATGTATATACAAACACACAGTACTTATTGATAAATTTGAATTT 2022
QY 702 IleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGln 721
Db 2023 CTGCCAATTAATCTCGTTCTATAAGAGAGGTAGAGAGAAACAAAAATTAGAAACAGTCAA 2082
QY 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db 2083 CAAATTAATTAATACATTTTATGCAATCTCTATAAAA 2118
RESULT 3
US-09-224-024-30
; Sequence 30, Application US/09224024
; Patent No. 6056953
; GENERAL INFORMATION:
; APPLICANT: Leslie Hickie
; APPLICANT: Jewel Payne
; TITLE OF INVENTION: Materials and Methods for the Control of
; TITLE OF INVENTION: Calliphoridae Pests
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/224,024
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/856,226
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA79
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2061 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-224-024-30
Alignment Scores: 1.61e-85 Length: 2061
Pred. No.:

Score:	899.50	Matches:	247
Percent Similarity:	48.32%	Conservative:	126
Best local Similarity:	31.9%	Mismatches:	270
Query Match:	23.25%	Indels:	129
DB:	3	Gaps:	30

US-10-783-417-2 (1-735) x US-09-224-024-30 (1-2061)

Qy	1	MetAsnGlnAsnAsnAspAsnAenGluTyrGluIleAspSerHisThrSerProTyr	20
Db	:	:::::	:::::
Db	1	ATGAATCCATATCAAAATAAGAATGAATGAATAATTCATAGCTCCCATCGTTT	60
Qy	21	PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro	40
Db	61	-----ACAAATCTAAATAACTATTCTAGATATCCATTAGCAAATAAGCAAATCAACCA	114
Qy	41	LeuGlnAsnThrAsnTyrLysGluTrpLeuAsnMetCysGlnGlyAsnThrGlnTyrGly	60
Db	115	CTGAATAAACACGAATTAACAAGATGGCTCANTGTGTCAAGATTAATCAACATATATGC	174
Qy	61	AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle	80
Db	175	AATAATGCCGGGAATTTGTTAGTTCAGAACTATTGTTGGAGTAGTCAGCGATTATAT	234
Qy	81	ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly	100
Db	235	GTAGTAGGAACATATCTTAGGA-----GCTTTTGCTGCCCT-----	270
Qy	101	IleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly	120
Db	271	GTCYTAGTCGCGGTATTAATATCTTTGGAGACTTTGTGCCGATCTTTGG---CAAGA	327
Qy	121	GluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThr	140
Db	328	TCTGACCCTGCAATGTTTGGCAGGATTTGTTAAACATCGGA-----GGAAGG	375
Qy	141	ProLeuThrGlu-----SerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPhe	157
Db	376	CCTATACAGAATAAGATAAAAACATAAATAATGACTAACTTCTATCCTAACACCTATA	435
Qy	158	ArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeu	177
Db	436	AAAAATCAACTTGATAATATCAAGAAATTTTCGATAAATGGAGCCACGACGT-----	489
Qy	178	GlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIle	197
Db	490	-----ACACACGCTAATGCTAAAGCAGTACATCATCTCTTACT	528
Qy	198	ArgPheGluAsnValHiAsn-----AspPheIleArgGluIleProGlyPheGln	214
Db	529	ACCTTAGAACCTATAATAGATAAAGATTTAGATATGTTAAAAATAATGCTAGCTATCGA	588
Qy	215	LeuGluThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAsnPheHisLeu	234
Db	589	ATACCAACA-----CTCCCTGTCATATGCACAAATAGTACTTTGGCACCTGG	633
Qy	235	AsnLeuLeuGlnGlnGlyValAlaLeuAlaAspGluTrpAsnAlaAspIleHisProSer	254
Db	634	AATTTATTAACATCGCTGCTACCTATTACAATATATGG-----CTGC AAAATCAA	684
Qy	255	GlnIleGluProAsnAlaGlyThrSerAspTyrTyrLys---LeuLeuLysGluAsn	273
Db	685	GGTATAAATCCAAGTATTTCAATTCATTAATTACTATCAGGGCTATTTAAAAACGTAAA	744
Qy	274	IleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArg	293
Db	745	ATACAAAGAATATACTGACTATTGTATACAAACGCTACAAATGCAGGAGCTACTATGATTGA	804
Qy	294	AspGluProAsnMetLysTrpSerIlePheAsnAspTyrTrpArgTyrMetThrIleThr	313
Db	805	ACTAATACTAACCGCAACATGGATATGATATATCTTACCGTTTGAATGATCTTAATCT	864
Qy	314	ValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIle	333

[illegible]

Db 1783 CCAATGCTACAGATTTAAACATATGCAGATTTTGGATATGATTAACATTTTCCAGAACAGATT 1842

Qy 673 -----ThrLeuProLeuAsnArgAsnIleProPheIle 683

Db 1843 CCAATAAACAACATTTGAAGGAGAAGACACTTTTATTAAATGACC-----TTA 1887

Qy 684 PheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspIlysrIleGluPheIlePro 703

Db 1888 TATGGTACACCAAAATCAATTCATATATAATA---TATATTGACAAAATCGAAATTTATTCCA 1944

Qy 704 IleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLys 723

Db 1945 ATCACTCAATCTGTATTAGATTATACAGAGAACCAAAATATAGAAAAACACAGAAAATA 2004

Qy 724 IleAsnThrPhePheThrAsnHisThrLysThrLeu 735

Db 2005 GTGAATGATTTATTCTGTTAAATTAACAAAGTCTTT 2040

RESULT 4

PCT-US94-07902-30

Sequence 30, Application PC/TUS9407902

GENERAL INFORMATION:

APPLICANT: Street address: 4980 Carroll Canyon Road

APPLICANT: City: San Diego California

APPLICANT: State/Province: California

APPLICANT: Country: US

APPLICANT: Postal code/Zip: 92121

APPLICANT: Phone number: (619) 453-8030 Fax number: (619) 453-6991

APPLICANT: Telex number:

TITLE OF INVENTION: Materials and Methods for the Control of

TITLE OF INVENTION: Calliphoridae Pests

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: David R. Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/07902

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Saliwanchik, David R.

REGISTRATION NUMBER: 31,794

REFERENCE/DOCKET NUMBER: MA79

TELECOMMUNICATION INFORMATION:

TELEPHONE: 904-375-8100

TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 2061 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

PCT-US94-07902-30

Alignment Scores:

Pred. No.: 1.61e-85 Length: 2061

Score: 899.50 Matches: 247

Percent Similarity: 48.32% Conservative: 126

Best Local Similarity: 31.99% Mismatches: 270

Query Match: 23.25% Indels: 129

DB: 6 Gaps: 30

US-10-783-417-2 (1-735) x PCT-US94-07902-30 (1-2061)

Qy 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20

Db 1 ATGAATCCATATCAAAAATAAGAAATGAATATGAAATATTCAATGTCTCCATCCAAATGGTTT 60

Qy 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40

Db 61 -----AGCAAGTCTAATAACTATTCTAGATATCCATTAGCAAAATAAGCCAAATCAACCA 114

Qy 41 LeuGlnAsnThrAsnTyrLysGluIleLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60

Db 115 CTGAATAACACAGAAATACAAAGATTGGCTCAATGTGTCTCAAGATAATCAACAATATGCG 174

Qy 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80

Db 175 AATAATGCGGGGAATTTTGTAGTTCTGAAACTATTGTGGAGTTAGTGCAGGTATTATT 234

Qy 81 ValSerGlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGly 100

Db 235 GTAGTAGGAACATATGTTAGGA-----GCTTTTGTCTGCCCT----- 270

Qy 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120

Db 271 GTCTTAGCTCAGGTATAATATCTTTTGGACTTTGTTCGGACTTTTGG-----CAAGGA 327

Qy 121 GluGlnAspIlyThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThr 140

Db 328 TCTGACCTCCAAATGTTTGGCAGGATTTGTTAAACATCGGA-----GGAAGG 375

Qy 141 ProLeuThrGlu-----SerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPhe 157

Db 376 CCTATACAGAAATAGATAAAACATAAATAATCTACTAACTTCTATCGTAACCACTATA 435

Qy 158 ArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeu 177

Db 436 AAAAAATCAACTTCATAAATATCAAGAAATTTTCGATAAATGGGAGCCAGCACGT----- 489

Qy 178 GlnAlaProGlyLeuProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIle 197

Db 490 -----ACACAGCTAATGCTAAAGCAGTACATGATCTCTTTACT 528

Qy 198 ArgPheGluAsnValHisAsn-----AspPheIleArgGluIleProGlyPheGln 214

Db 529 ACCTTAGAACCTATATAAGATAAAGATTAGATATGTTAAAAATAAATAGCTAGCTATCGA 588

Qy 215 LeuGluThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeu 234

Db 589 ATACCAACA-----CTCCCTGCATATGCACAAATAGCTACTTGGCAGCTTG 633

Qy 235 AsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSer 254

Db 634 AATTTATAAACATGCTGCTACCTATTACATAATATGG-----CTCAGAAATCAA 684

Qy 255 GlnIleGluProAsnAlaGlyThrSerAspTyrTyrLys-----LeuLeuLysGluAsn 273

Db 685 GGTATAAATCCAAAGTACTTTTCAATTCATCTAATTAATCTAGGCTATTTTAAACGTA 744

Qy 274 IleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArg 293

Db 745 ATACAAAGATATATCTGACTATTGTATACAAACGTCACAAATGCAGGACTAACTATGATTGA 804

Qy 294 AspGluProAsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIleThr 313

Db 805 ACTAATACTAACCAACATGGAATATGTAATACTTACCCTTTAGAAATGACTCTAACT 864

Qy 314 ValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIle 333

Db 865 GTGTAGATCTTATGCTATTTTCCAAATTAATGACCCAGAAATAAT----- 912

Qy 334 GlyGlyIleGluValLysGlyIleLysAsnGlnLeuThrArgGluIleTyrThrGlu 353

Db 913 -----CCAAATAGGAGTTAAATCTGAACCTTACCAGAGAGTTTAT---ACGAAT 957

Pred. No.: 6-55e-72 Length: 3507
 Score: 775.50 Matches: 217
 Percent Similarity: 46.39% Conservative: 136
 Best Local Similarity: 28.52% Mismatches: 289
 Query Match: 20.04% Indels: 119
 DB: 2 Gaps: 27

US-10-783-417-2 (1-735) x US-08-315-468-3 (1-3507)

QY 5 AsnAspAsnGluTyrGluIleAspSerHisThrSerProTyrPheProAsnArg 24
 DB 10 AATAATCAAAATGAATATGAATATATAGATGCGACACTTCT-----ACATCT 57
 QY 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
 DB 58 GTATCAATGATCTTAACAGATACCTTTTTCGGAATGAGCCCAAAATGCGGTACAAAT 117
 QY 44 ThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
 DB 118 ATGGATTATAAGATATTATTAATAATGTCTCGGGAATCTTAGTGAATACCTCGGTCA 177
 QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSer 82
 DB 178 CCTGAGTATTTCTAAGCGAGCAAGAT-----GCAGTTAAGCGCGCAATTGATAGTA 231
 QY 83 GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIle 102
 DB 232 GGTAAATTTACTAACAGGTTTAGGG-----GTTCCATTGTT 267
 QY 103 GlyAlaIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGlyGluGln 122
 DB 268 GGGCCGATAGTAGTCTTTACTCAACTTATTATTTCTGTGGCCTTCA-----AAA 321
 QY 123 AspLysThrValTyrThrGlnPheIleLysMetGlyLysIlePheValAspThrProLeu 142
 DB 322 CAAAGAGTCATGGGAATTTTATGGAAACAGTAGAAGCACTCATTAATCAAAATA 381
 QY 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
 DB 382 GCAGATATGCAAGGAATAAGCGCTTCGGAATTTGGAAGCGCTAGGGAATAATTACCA 441
 QY 163 SerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
 DB 442 TTATATCTAACTCGCTTGAAGAGTGAAGAA-----474
 QY 183 ProProSerSerAlaLeuGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal 202
 DB 475 AATCCAAATGGTTCA-----AGAGCCTTACGAGATGTTTGAATCGATTGAAATCCTG 528
 QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
 DB 529 GATAGTTTATTACGCAATATATGCAATCTTTTCGAGTGAACAATTTTGAAGTACCATT 588
 QY 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlyAlaGlu 242
 DB 589 CTTACAGATATACAATGGCAGCAACCTACATTTACTTTTATTAAAGGCGCATCAATT 648
 QY 243 LeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
 DB 649 TTTGGAGAAGATGGGGATTG-----TCTACAAGCACT 681
 QY 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
 DB 682 ATTAATAACTACTATTAATCTGCAATGAAATCTTATCGCAGAAATTTCTGACCCTGTGTA 741
 QY 283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTyrSerIle 302
 DB 742 AAGTGTATGAATCGTTTAGCAAAATTTAAAGGCTCGAGCGCTAAACAATGGATTGAC 801
 QY 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
 DB 802 TATAACCAATTCGTAGAGAATGACATTCACGCGTTAGACGTTGTTGATATTATTTC 861

QY 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIleLys 342
 DB 862 AACTATGATACGCGTACGTAT-----CCACTGGCAACAACA 897
 QY 343 AsnGluLeuThrArgGluIleTyrThrGluIleAsnPheAspArgLeuProGlnLeu 362
 DB 898 GCTCAGCTTACAGGGAAGTATATACAGATCCACTTGGCGGGTAGATGTCCTTAATATT 957
 QY 363 ArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPheLys--- 381
 DB 958 GGCTCTCGTATGACAAAGCACCTCTTCTTCAGAAATAGAAAAAGCGCTATTCTGCTCA 1017
 QY 382 -----LeuPheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsnPhe--- 398
 DB 1018 CCTCATGTGTTGATTATATAACGGGACTCACAGTTTATACAAAAAACAGTAGCTTCACT 1077
 QY 399 -----GlyAsnArgLeuValGlyIleSerAsnArgAsp 409
 DB 1078 TCTGATCGTTATATGAGATATATGGCTGGTCATCA-----ATAAGCTATATAGCAT 1128
 QY 410 AlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySerPro 429
 DB 1129 ATCGGTACG---AGTAGTACCTTTACACAG---ATGTATGGAACCAATCAAAATTACAA 1182
 QY 430 ThrThrLysThrIleArgProPheGluSerTyrLysVal-----442
 DB 1183 AGTACTAGCAATTTT---GATTTTACGAATTTACGATATTTTACAAGACTTTTATCAAAATGGT 1239
 QY 443 SerIleValThrAspArgGlnSerProValSer-----ProIle 456
 DB 1240 GCAGTACTCTTGATATAGTTTACCTCGGTATACGTATACATTTTGGATGCGCAGAA 1299
 QY 457 GlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsnThr 476
 DB 1300 ACCGAGTTTTTATGGTAAATCAATTG-----ATAATATACC 1335
 QY 477 LeuLysTyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPhe 496
 DB 1336 AGAAG-----ACGTAAACGTATATAAACCCAGCTTCCAAAGATATTATA 1377
 QY 497 ProArgLysAspCysAsnLeuValIleAspProGlyCysSer-----ProAsnPhe 514
 DB 1378 GATCGGACAGAGATTCGGAATTAGAATTCCTCCAGAACTTCAGGTCAACCAAAATTAC 1437
 QY 515 AsnAsnTyrSerHisIleLeuSerHis-----PheSerLeuPheThrTyrSer 530
 DB 1438 GAGTCATATAGCCATAGATTAGGTCAATTATACATTTATTACTCCAGTTCACACTAGCAG 1497
 QY 531 TyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSer 550
 DB 1498 TATGTA-----CCTGATTTTCTTGGACACATCGG 1527
 QY 551 SerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLys 570
 DB 1528 AGTGCAGATCTAAACAATACAGTTTAAAGTGGCGAAATCAACCAAAATACCGGGGCGAAG 1587
 QY 571 GlyAsnAsnLeuAspThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsn 590
 DB 1588 TCTAGCACCATGAGCAGAAATATATATAAAGGCGGTGTTATACAGGGGGAGAC 1647
 QY 591 LeuValTyrLeuGlnSerGln---GlyArgLeuGluIleThrCysGluThrProAsnSer 609
 DB 1648 TTAGTGGCTTAAACGGACCGCATCGGAAGTTGTGAGTTTCAGATGATCTTTCCAGAGTCT 1707
 QY 610 ThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuPro 629
 DB 1708 ---CAACGATTCGCTATTCGATTCGTTTCTAATGAACACTAGT-----1752
 QY 630 AsnIleSerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsnAsnThr 649
 DB 1753 -----TATATTAGTTTATACGGACTAAACCAAAAGCGGAACCTTTAAATTCACACAGACA 1806
 QY 650 PheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPhePro 669

Db 1807 TATTCTAATAAATGAATGATTTTACATATATGATTTCAATATATAGATATCCA 1866
QY 670 SerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspVal 689
Db 1867 AGAGTCATTTTCAGTAAATGCTTCTTCAACATACACAGAGTTTATCTATAGGTATACAAACG 1926
QY 690 SerAsnSerIleLeuIleAspLysIleGluPheIleProIleThrSerSerMetHis 709
Db 1927 AATACAAATTTATTTTATAGCCGAATCGAATTCATCCAGTAGATGAGACATATGAA 1986
QY 710 GluAsnArgGluLysGlnLysLeuGluThrIleGlnThrIleAsnThrPhePheThr 729
Db 1987 GCGGAAACGGAT-----TTAGAGCGCAAGAAAGCAGTGAATGCCCTGTTTACG 2037
QY 730 Asn 730
Db 2038 AAT 2040

RESULT 6
US-08-315-468-5
; Sequence 5, Application US/08315468
; Patent No. 5554534
; GENERAL INFORMATION:
; APPLICANT: Michaels, Tracy Ellis
; APPLICANT: Fonceerrada, Luis
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Process for Controlling Scarab Pests
; TITLE OF INVENTION: with Bacillus thuringiensis Isolates
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/315,468
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,941
; FILING DATE: 01 FEB 1993
; APPLICATION NUMBER: 07/828,430
; FILING DATE: 30-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/808,316
; FILING DATE: 16-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA73.C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1953 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: tolworthi

INDIVIDUAL ISOLATES: 43F
IMMEDIATE SOURCE: XLI-Blue (pMI,98-4), NRRL B-18291
CLONE: E. coli
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1953
US-08-315-468-5
Alignment Scores:
Pred. No.: 1.05e-61 Length: 1953
Score: 676.00 Matches: 209
Percent Similarity: 43.76% Conservative: 117
Best Local Similarity: 28.05% Mismatches: 283
Query Match: 17.47% Indels: 136
DB: 2 Gaps: 29
US-10-783-417-2 (1-735) x US-08-315-468-5 (1-1953)
QY 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 1 ATGAATCCAAACAAT---CGAAGTGAATATGATACGNTAAAGGTTTACACCTAACAGTGAA 57
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 58 TTGCA-----ACTAACCATATCATATCTTTAGCTGACAAATCCAATTCGACA 108
QY 41 LeuGlnAsnThrAsnTyrLysGluTrpLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
Db 109 CTAGAAGAATTAATATAAAGAAATTTTAAAGATGACTGCAGACAAATTCACG----- 162
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80
Db 163 -----GAAGTGCTAGACAGCTCTACAGTAAAGATGCAGTTGGGACAGCAATTTCT 213
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
Db 214 GTTGAGCAGATTTAGGTGTGTAGGG-----GTCCA 249
QY 101 IleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120
Db 250 TTTGCTGGGGCGCTCACCTTCATTTTATCAATTCCTTAACGCTATATGGCAAGT--- 306
QY 121 GluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db 307 -----GATGCTGACCCATGGAAGCTTTTATGGCACAAGTGGCAAGTACTGATAGATAAG 360
QY 141 ProLeuThrGluSerIleLysGlnLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
Db 361 AAAATAGAGGAGTATGCTTAAAGTAAAGCTCTTCGAGAGTTACAGGGCTTCAAAATAAT 420
QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaPro 180
Db 421 TTTGAAGATTATGTAATGCGTTGGATTCTCTGGAAGAAA----- 459
QY 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGlu 200
Db 460 ---GCGCCTGTAATTTACGAAGTCCGAAGAACCAAGATCGAATAAGAGAACTTTTTTCT 516
QY 201 AsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThr 220
Db 517 CAAGCAGAAAGCCATTTTCCATGCGCTCATTTGCGGTTTCCAAATTCGAAGTT 576
QY 221 LeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGly 240
Db 577 CTGTTTCTACCAACATATGCAAGCTGCAATATACACATTTATGCTATTAAAGATGCT 636
QY 241 AlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAla 260
Db 637 CAAGTTTTTGGAGAAGATGGGATAT----- 663
QY 261 GlyThrSerAspAsp-----TyrTyrLysLeuLeuLysGluAsnIleProLysTyr 277
Db 664 ---TCTTCAGAAGATATTGCTGAATTTTATCAAGACAAATTAAGAACTTACGCAACAATAC 720

Qy	613	TyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSer	632
Db	1696	TATCGCGTAAAGATACGCTATGCTTCAACC-----ACTAACCTACGA	1738
Qy	633	LeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnThrPheSerGly	652
Db	1738	CTTTTCGTG-----CAAAATCAAAACAATGATTTTCTTGTC	1773
Qy	653	ThrAsnTyrAsn-----AsnLeuGlnTyrGlyAspPheGlyTyr-----PheGln	667
Db	1774	ATCTACATTAAATAAACAATATGATATAGAT-----GGTGATTTAACATATCAAAACATTGTAT	1830
Qy	668	PheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAla	687
Db	1831	TTCCCAACTAGT-----AATTCCTAATATGGGATTCCTGGTGATACAAAT	1875
Qy	688	Asp-----ValSerAsnSerIleLeuIleAspLysIle	699
Db	1876	GACTTTATATAGACAGACAGAACTCTTCGTTTCTAATGAAAAAATCTATATAGATAAGATA	1935
Qy	700	GlupheIleProIle 704	
Db	1936	GAATTTATCCAGTA 1950	
RESULT 7			
	US-07-941-650A-3		
	; Sequence 3, Application US/07941650A		
	; Patent No. 6294184		
	; GENERAL INFORMATION:		
	; APPLICANT: Uyeda, Kendrick A.		
	; APPLICANT: Bradfisch, Gregory A.		
	; TITLE OF INVENTION: Process for Controlling Lepidopteran Pests		
	; NUMBER OF SEQUENCES: 7		
	; CORRESPONDENCE ADDRESS:		
	; ADDRESSEE: David R. Saliwanchik		
	; STREET: 2421 N.W. 41st Street, Suite A-1		
	; CITY: Gainesville		
	; STATE: FL		
	; COUNTRY: USA		
	; ZIP: 32606		
	; COMPUTER READABLE FORM:		
	; MEDIUM TYPE: Floppy disk		
	; COMPUTER: IBM PC compatible		
	; OPERATING SYSTEM: PC-DOS/MS-DOS		
	; SOFTWARE: PatentIn Release #1.0, Version #1.25		
	; CURRENT APPLICATION DATA:		
	; APPLICATION NUMBER: US/07/941,650A		
	; FILING DATE: 19920908		
	; CLASSIFICATION: 435		
	; PRIOR APPLICATION DATA:		
	; APPLICATION NUMBER: US 07/758,020		
	; FILING DATE: 12-SEP-1991		
	; PRIOR APPLICATION DATA:		
	; APPLICATION NUMBER: US 07/658,935		
	; FILING DATE: 21-FEB-1991		
	; PRIOR APPLICATION DATA:		
	; APPLICATION NUMBER: US 07/642,112		
	; FILING DATE: 16-JAN-1991		
	; ATTORNEY/AGENT INFORMATION:		
	; NAME: Saliwanchik, David R.		
	; REGISTRATION NUMBER: 31,794		
	; TELECOMMUNICATION INFORMATION:		
	; TELEPHONE: 904-375-8100		
	; TELEFAX: 904-372-5800		
	; INFORMATION FOR SEQ ID NO: 3:		
	; SEQUENCE CHARACTERISTICS:		
	; LENGTH: 1953 base pairs		
	; TYPE: nucleic acid		
	; STRANDEDNESS: double		
	; TOPOLOGY: linear		
	; MOLECULE TYPE: DNA (genomic)		
	; HYPOTHETICAL: NO		
	; ANTI-SENSE: NO		

ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: toleworthi
; INDIVIDUAL ISOLATE: 43F
; IMMEDIATE SOURCE:
; CLONE: E. coli XL1-Blue (pM1.98-4), NRRL B-18291
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1953
; US-07-941-650A-3

Alignment Scores:
Pred. No.: 1.05e-61 Length: 1953
Score: 676.00 Matches: 209
Percent Similarity: 43.76% Conservative: 117
Best Local Similarity: 28.05% Mismatches: 283
Query Match: 17.47% Indels: 136
DB: 3 Gaps: 29

US-10-783-417-2 (1-735) x US-07-941-650A-3 (1-1953)

QY 1 MetAsnGlnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 1 ATGAATCCAAACAAT---CGAAGTGAATATGATACGATAAAGGTTACACCTTAACAGTGAA 57
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 58 TTGCCA-----ACTAACCAATAATCAATATATCTTTAGCTGACATCCAAATTCGACA 108
QY 41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
Db 109 CTAGAGAATAAATTAAGAATTTTAAAGATGATCTGAGATGCTGACACAATTCACG----- 162
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80
Db 163 -----GAAGTGCTAGACAGCTCTACAGTAAAGATGCAGTTCGGACAGGAAATTCCT 213
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
Db 214 GTTGTAGGACAGATTTTAGGTGTGTAGGG-----GTTCCA 249
QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGly 120
Db 250 TTTGCTGGGCGCTCACTTCATTTCAATCATTTCTTAAACGCTATATGGCCAAAGT--- 306
QY 121 GluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db 307 -----GATGCTGACCCATGGAAGGCTTTTATGGCACAAAGTGGAGTACTGTAGATAAG 360
QY 141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
Db 361 AAAATAGAGGAGTATGCTAAAGTAAAGCTCTTCAGAGTTACAGGCTCTTCAAAATAAT 420
QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaPro 180
Db 421 TTGAAGATTTATGAATGCGTTGGATTCCTCGAAGAAA----- 459
QY 181 GlyLeuProProSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGlu 200
Db 460 ---GCGCTGTAATTTACGAAGTCGAAGAACCCAGATCGAATAAGAGAACATTTTCTCT 516
QY 201 AsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThr 220
Db 517 CAAGCAGAAAGCCATTTTCGTAATTCATCGCGTCATTTGCGGTTTCCAAATTCGAAGTT 576
QY 221 LeuLeuLeuProTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGly 240
Db 577 CTGTTTCTACCAATATGACAAAGTGCACAAATACACATTATTATGCTATTAAAGATGCT 636
QY 241 AlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAla 260
Db 637 CAAGTTTTTGGAGAAGAATGGGATAT----- 663

QY 261 GlyThrSerAspAsp-----TyrTyrLysLeuLeuLysGluAsnIleProLysTyr 277
Db 664 ---TCTTCAGAAGATATTGCTGAATTTTATCAAGACAAATTAATAACTTACGCAACAATAC 720
QY 278 SerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsn 297
Db 721 ACTGACCATGCTGCTCAATATGATGATTAATGATGATTAATGATTAAGAGGTTCACTTAT 780
QY 298 MetLysTyrSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThr 317
Db 781 GATGATGGGTCAAAATTTAAACCGTTTTCGACAGAAATGACATTAACTGATTATAGATCTA 840
QY 318 IleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGlu 337
Db 841 ATTGTATTATTCCTATTTATGATGTTGCGGTATACTCA----- 879
QY 338 ValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrGluIleAsnPheAsp 357
Db 880 ---AAGGAGTTAAACAGAACTAAACAGACATTTTACAGATCCAATT---TTTACA 933
QY 358 ArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArg 377
Db 934 CTCAATGCTCTTCAAGAGTATGACCAACTTTTTCGAGTATAGAAACTCTATTTCGAAAA 993
QY 378 AlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsn 397
Db 994 CCT-----CATTTTATTTGATTTATTTGCGTGGGATTTGAATTTTCATACG----- 1035
QY 398 PheGlyAsnArgLeu---ValGlyIleSerAsnArgAspAlaProThrTyr----- 413
Db 1036 -----CCTCTTCGACCTGGTTACTCTGGGAAAGATTCCTTCAATTTATTTGGTCTGGT 1086
QY 414 -----SerAsnThrIleThrGluThrLeuTyr 422
Db 1087 AATTATGTAGAACTAGACCTAGTATAGGATCTAATGATACATCACTCCTCCCATTTTAT 1146
QY 423 GlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSer----- 439
Db 1147 GGAGATAAATCTATTGAACCTATACAAAAGCTA-----AGCTTTGTGCGACAAAAGTT 1200
QY 440 TyrLysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHis 459
Db 1201 TATCGAACTATAGCTAATACAGACATAGCGGCTTTTCGGATGGCAAGATA-----TAT 1254
QY 460 PheIleIleAsnGlnIleGluLeu---TyrLeuAsnGlySerSerAsnThrLeuLys 478
Db 1255 TTTGGTGTACGAAAGTTGATTTTAGTCAATATGATGATCAAAAAATGAACT----- 1308
QY 479 TyrSerAlaGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArg 498
Db 1309 -----AGTACACAAACATATGATTCAAAAAGATCAATAGGCTATTTAGGT 1353
QY 499 LysLysAspCysAsnLeuValIleAspProGlyCysSer-----ProAsnPheAsnAsn 516
Db 1354 GCACAGGATCTATCGACCAATTTACCACAGAAACACAGATGAACCACTTGAAAAAGCA 1413
QY 517 TyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGln 536
Db 1414 TATAGTCATAGCTTAATAGCGAGAATGTTCTTA-----ATGCGAG 1455
QY 537 LeuGlnIleLeuAspThrGlyValLeuGlyTyrHisSerSerValAspArgTyrAsn 556
Db 1456 GACCGTGTGGAACCAATTCCTATTTTACTTTGGACACATAGAAAGTGTAGACTTTTATAT 1515
QY 557 AlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThr 576
Db 1516 ACAATTGATGCTGAAAAAATTTACTCAACTTCAGTAGTGAAGACATATGCTTGTCTTCA 1575
QY 577 AsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln--- 595
Db 1576 GGCGCTTCCCATTTATGAAGGTCAGGATTCACAGAGAAATTTACTATTCTTCAAAAGAA 1635
QY 596 -----SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSer 612

```
Db 1636 TCTAGTAATTCATGCTAAATTTAAAGTTACCTTAAATTCAGCAGCCTGTTTACACGA 1695
Qy 613 TyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSer 632
Db 1696 TATCGCGTAAGATAACGCTACTGCTCAACC-----ACTAACCTACGA 1737
Qy 633 LeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGly 652
Db 1738 CTTTTCGTG-----CAAAATTCNAACAATGATTTTCTGTG 1773
Qy 653 ThrAsnTyrAsn-----AsnLeuGlnTyrGlyAspPheGlyTyr-----PheGln 667
Db 1774 ACTACATTAATAAAACTATGATATAGAT---GGTGAATTAACATATCAACATTTGAT 1830
Qy 668 PheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAla 687
Db 1831 TTGCGCACTAGT-----AATTCATATATGGGATTCCTCGGTGATACAAAT 1875
Qy 688 Asp-----ValSerAsnSerIleLeuIleAspLysIle 699
Db 1876 GACCTTTATATAGCAGCAGAACTCTTCGTTCTTAATGAAAAAATCTATATAGATAAGATA 1935
Qy 700 GluPheIleProfile 704
Db 1936 GAAATTTATCCAGTA 1950

RESULT 8
PCT-US92-00040-1
; Sequence 1. Application PC/TUS9200040
; GENERAL INFORMATION:
; APPLICANT: Donovan, William P.
; APPLICANT: Ruper, Mark J.
; APPLICANT: Slaney, Annette C.
; TITLE OF INVENTION: BACILLUS THURINGIENSIS cryIIIC(b) TOXIN
; TITLE OF INVENTION: GENE AND PROTEIN TOXIC TO COLEOPTERAN INSECTS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
; ADDRESSEE: Nadel
; STREET: 1601 Market Street, 36th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/00040
; FILING DATE: 19920103
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/649,562
; FILING DATE: 31-JAN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Egolf, Christopher
; REGISTRATION NUMBER: 27633
; REFERENCE/DOCKET NUMBER: 7205-29 P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-757-1590
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2430 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 144..2099
```

```
PCT-US92-00040-1
Alignment Scores:
Pred. No.: 3,24e-60 Length: 2430
Score: 663.50 Matches: 217
Percent Similarity: 41.87% Conservative: 115
Best Local Similarity: 27.36% Mismatches: 300
Query Match: 17.15% Indels: 161
DB: 6 Gaps: 31

US-10-783-417-2 (1-735) x PCT-US92-00040-1 (1-2430)
Qy 1 MetAsnGlnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 144 ATGAAGTCAACCAAT---CGAAGTGAACATGATACGATAAGGTTACACCTTAACAGTAA 200
Qy 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 201 TTGCCA-----ACTAACCATATCAATATTCCTTTAGCTGACAATCCAAATTCGACA 251
Qy 41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGln---GlyAsnThrGlnTyr 59
Db 252 CTAGAAGAATTAATATAAAGAAATTTTAAAGATGACTGAAGACAGCTTCTACCGAAGTG 311
Qy 60 GlyAspAsnPheGluThrPheAlaSerAlaAspThrIleAla---AlaValSerAlaGly 78
Db 312 CTAGACAAC-----TCTACAGTAAAGATGCGATGTTGGGACAGGA 350
Qy 79 ThrIleValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyPro 98
Db 351 ATTTCTGTGTGAGGCGAGATTTTAGGTGTGTAGGA----- 386
Qy 99 IleGlyIleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrPro 118
Db 387 GTTCCATTTGCTGGGCACTCACTTCATTTTATCAATCATTTCTTGACACTATATGCGCA 446
Qy 119 AlaGlyGluGlnAspLysThrValThrGlnPheIleLysMetGlyGluIlePheVal 138
Db 447 AGT-----GATGCTGACCCATCGAAGGCTTTTATGGCACAAGTTGAAGTACTGATA 497
Qy 139 AspThrProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArg 158
Db 498 GATAAGAAAATAGAGAGATATGCTAAAGTAAAGCTCTTCGAGAGTTACAGGGTCTTCAA 557
Qy 159 GlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGln 178
Db 558 AATAATTTCGAAGATTAATGTTAATGCGTTAAATTCCTGGAGAGAAA----- 602
Qy 179 AlaProGlyLeuProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArg 198
Db 603 -----ACACCTTTAAGTTTTCGGAAGTAAAGAACCAAGATCGAATAAGGGAACCTT 653
Qy 199 PheGluAsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyr 218
Db 654 TTTTCTCAAGCAGAAAAGTCAATTTTCGTTAAATTCATGCGCGTCATTTCAGTTTCCAAATTC 713
Qy 219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAsnPheHisLeuAsnLeuLeuGln 238
Db 714 GAAAGTGTCTGTTTCTACCAACATATGCAACAGCTGCAAAATACACATTTATTGCTATTAAA 773
Qy 239 GlnGlyAlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluPro 258
Db 774 GATGCTCAAGTTTTTGGAGAGAAATGGGATAATTCT----- 809
Qy 259 AsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSer 278
Db 810 TCAGAAAGATGTTGCTGAATTTTATCATAGACAATTTAAACTTACG---CAACAATACACT 866
Qy 279 AsnTyrCysAlaAsnThrTyrArgThrGlyLysAsnLeuArgAspGluProAsnMet 298
Db 867 GACCATTTGTCAAATGGGTATAATGTTGGATTAATGTTTAAGAGGTTCAACTATGAT 926
Qy 299 LysTrpSerIlePheAsnAspTyrArgArgTyrMetThrThrIleThrValLeuAspThrIle 318
```

Db 927 GCATGGGTCAAATTAACCGTTTTCGAGAGAAATGACTTTAACTGATTAGATCTAATT 986
Qy 319 SerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyLeuVal 338
Db 987 GTACTTTTCCATTTTATGATGTTTCGGTTATACTCA 1022
Qy 339 LysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArg 358
Db 1023 AAGGGTGTAAACAGAACTAACAGAGACATTTTACGGATCCAATT 1079
Qy 359 LeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAla 378
Db 1080 AATACTCTTCAGGAGTATGACCAACTTTTGTAGTATAGAAACTCTATTTCGAAACCT 1139
Qy 379 SerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsnPhe 398
Db 1140 CATTATTTGATTATTTACAGGATTTTACAGGATTTGAAATTTTCATACG 1178
Qy 399 GlyAsnArgLeuValGlyIleSerAsnArgAspAlaProThrTyr 413
Db 1179 CGTCTTCACTGTTACTCTGCGAAGATTTCTTCAATTATGCTGGTAAAT 1232
Qy 414 SerAsnThrIleThrGluThrLeuTyrGly 423
Db 1233 TATGTAGAACTAGACCTAGTATAGGATCTAGTAAGACAAATTACTTCCCACTTTATGGA 1282
Qy 424 GluArgThrGlySerProThrThrThrIleArgProPheGluSer 440
Db 1293 GATAAATCTACTGAACCTGTACAAAAGTTA 1346
Qy 441 LysValSerIleValThrAspArgGlnSerProValSerProIleGlnProHisPhe 460
Db 1347 CGAACTATAGCTAATACAGAGCTAGCGGCTTGGCCGAATGCGCAAGATA 1400
Qy 461 IleIleAsnGlnIleGluLeu 475
Db 1401 GGTGTACGAAAGTTGATTTTAGTCAATATGATGATCAAAAATGAAACTAGTACACAA 1460
Qy 476 ThrLeuLysTyrSerAlaGlySerLeuSerAsnTyrGlnAsnThrThrPheGln 495
Db 1461 ACATATGATTTCAAAAAGAAACAATGGCCATGAGTGCACAGGATTTCTATTGACCAATTA 1520
Qy 496 PheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsn 515
Db 1521 CCACAGAAACAACAGAT 1553
Qy 516 AsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeu 535
Db 1554 GCATATAGTATCAGCTTAATACGGGAATGTTCTTA 1595
Qy 536 GlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyr 555
Db 1596 CAGGACCGTGTGGAACAATTCATTTTACTTGGACATAGAGTGTAGACTTTT 1655
Qy 556 AsnAlaIleSerAspLysIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp 575
Db 1656 AATACAAATGATCTGAAAAGATTACTCAACTTCCAGTAGTGAAGCATATGCTTGTCT 1715
Qy 576 ThrAsnSerLysValIleGluGlyProGlyHisThrGlyLysLeuValTyrLeuGln 595
Db 1716 TCAGGTGCTTCCATTTATGAAGTCCAGGATTCACAGGAGAAATTTACTATTCTCTAAAA 1775
Qy 596 SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGln 611
Db 1776 GAATCTAGTAATCAATGCTAAATTTAAAGTTATACATTAATTCAGAGCTTGTACAA 1835
Qy 612 SerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIle 631
Db 1836 CGATATCGTATAGATACGCTATGCTTCTACC 1877
Qy 632 SerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsnThrPheSer 651

1878 CGACTTTTGTG-----CAAAATTCAAACAATGATTTTATT 1913
Qy 652 GlyThrAsnTyrAsn-----AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPhe 668
Db 1914 GTCATCTACATTAATAAAACTATGAATATAGAT-----GATGATTTAAACATATCAACATTT 1970
Qy 669 ProSerThrValThrLeuProLeuAsnArgAsnIleProPhe----- 682
Db 1971 GATCTCGCACTACT-----AATCTAATATGGGTTCTCGGGTGATACGAATGAA 2021
Qy 683 -----IlePheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIleGlu 700
Db 2022 CTTATAATAGGAGCAGATCTTTCGTTTCTAATGAAAATCTATATAGATAGATAGAA 2081
Qy 701 PheIleProIle----- 704
Db 2082 TTTATCCAGTACAAATTTGTAAGAGATTTTGAATGTAGGCGCATGCTCAAAATGAAAGA 2141
Qy 705 -----ThrSerSerMetHisGlnAsnArgGluLys 2141
Db 2142 ATAGGAAGGTGAATTTTTCATGTTAGGAAAGATTTTAAAGAAAAGCAACATGGAAG 2201
Qy 715 GlnLysLeuGluThrIleGlnThrLys-----IleAsnThr 726
Db 2202 TATACAGTACAAATATAGAAATAAAATTTATTAAACACA 2240

RESULT 9
US-08-996-441B-37
; Sequence 37, Application US/08996441B
; Patent No. 6023013
; GENERAL INFORMATION:
; APPLICANT: English, Leigh H.
; APPLICANT: Brussock, Susan M.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Bryson, James W.
; APPLICANT: Kulesza, Caroline A.
; APPLICANT: Walters, Frederick S.
; APPLICANT: Slatin, Stephen L.
; APPLICANT: Von Tersch, Michael A.
; APPLICANT: Romano, Charles
; TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996.441B
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: MECO:151
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1959 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS

: LOCATION: 1..1956
 US-08-996-441B-37

Alignment Scores:
 Pred. No.: 2,596-60 Length: 1959
 Score: 663.00 Matches: 208
 Percent Similarity: 43.51% Conservative: 114
 Best Local Similarity: 28.11% Mismatches: 292
 Query Match: 17.14% Indels: 126
 DB: 3 Gaps: 27

US-10-783-417-2 (1-735) x US-08-996-441B-37 (1-1959)

Qy	1	MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr	20
Db	1	ATGAATCCAAACAAT--CGAAGTGAACATGAT-----ACGATAAAGGTT	42
Qy	21	PheProAsnArg-----AsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsn	38
Db	43	ACACCTAACAGTGAATTCGAACCTAACCATATCAATATCCTTTAGCTGACAATCCAAAT	102
Qy	39	GlnProLeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGln---GlyAsnThr	57
Db	103	TCAACACTAGAAGAATTAATATTAAGAATTTTAAAGATGACTGAACAGCTTCTACG	162
Qy	58	GlnTyrGlyAspAsnPheGluThrPheAlaSerAlaAspThrIleAla---AlaValSer	76
Db	163	GAAGTGCTAGACAAC-----TCTACAGTAAAGATGCAGTTGGG	201
Qy	77	AlaGlyThrIleValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSer	96
Db	202	ACAGGAATTTCTGTTGTAGGCGAGATGCTTGGTGTGTAGGA-----	243
Qy	97	GlyProIleGlyIleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPhe	116
Db	244	-----GTTCCATTGCTGGGGCACTCACTTCATTTTATCAATCATTTCTTAACACTATA	297
Qy	117	TyrProAlaGlyGluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIle	136
Db	298	TGGCCAGT-----GATGCTCACCNTGGAGGCTTTTATGGCACAAGTTGAGTA	348
Qy	137	PheValAspThrProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGly	156
Db	349	CTGATAGATAAGAAAATAGAGGAGTATGCTAAAGTAAAGCTCTTGACAGTTACAGGCT	408
Qy	157	PheArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrPargLysLeuLysArg	176
Db	409	CTTCAAAATAATTTCGAAGATATATGTTAATGCGTTTAAATTCCTCGAAGAAA	459
Qy	177	LeuGlnAlaProGlyLeuProProSerAlaLeuGlnGlnAlaAlaLeuThrLeuLys	196
Db	460	-----ACACCTTTAAGTTGGCAAGTAAAGAACCAAGTCGATGAATAGG	504
Qy	197	IleArgPheGluAsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGlu	216
Db	505	GAACCTTTTCTCAAGCAGAGAAGTCATTTTCGTAATTCATCCGCTCATTTGCAGTTTCC	564
Qy	217	ThrTyrLysThrLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeu	236
Db	565	AAATTCGAAGTGCCTTTCTACCAACATATGCACAAGCTGCAAAATACACATTTATTGCTA	624
Qy	237	LeuGlnGlnGlyAlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIle	256
Db	625	TTAAAGATGCTCAAGTTTTTGGAGAGAGATGGGATATTCT-----	666
Qy	257	GluProAsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLys	276
Db	667	-----TCAGAAGATGTTGCTGAAATTTTATCATAGACAATTAACACTTACA---CAACAA	717
Qy	277	TyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro	296
Db	718	TACACTGACCATGTTGTTAATTTGGTATTAATGTTGATTAATTTAGAGAGTTTCAACT	777

QY 39 GlnProLeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGln---GlyAsnThr 57
Db 103 TCACACTAGAGAAATTAATTAAGAAATTTTAAAGATGACTGAAGACAGTCTACG 162
QY 58 GlnTyrGlyAspAsnPheGluThrPheAlaSerAlaAspThrIleAla---AlaValSer 76
Db 163 GAAGTGTCTAGACAAAC-----TCTACAGTAAAGATGCAGTGTGGG 201
QY 77 AlaGlyThrIleValSerGlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSer 96
Db 202 ACAGGAATTTCTGTGTAGGCGAGATTTAGGTGTGTAGGA----- 243
QY 97 GlyProIleGlyIleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPhe 116
Db 244 -----GTTCCATTGCTGGGCACTCACCTCATTTTATCATCTTTTAAACACTATA 297
QY 117 TrpProAlaGlyGluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIle 136
Db 298 TGGCCCAAGT-----GATGCTGACCCATGGAAGGCTTTTATGGCACAAGTTGNA GTA 348
QY 137 PheValAspThrProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGly 156
Db 349 CTGATAGATAAGAAAAATAGAGAGATATGCTAAAGATAAGCTCTTGACAGAGTTACAGGGT 408
QY 157 PheArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArg 176
Db 409 CTTCAAAATAATTTTCGGAAGATATGTTAAATGCGTTAAATTCCTGGAAGAAA----- 459
QY 177 LeuGlnAlaProGlyLeuProSerSerAlaLeuGlnGlnAlaLeuThrLeuLys 196
Db 460 -----ACACCTTTAAGTTTGGGAAGTAAAGAACCCAGATCGNATGAAG 504
QY 197 IleArgPheGluAsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGlu 216
Db 505 GAACCTTTTCTCAAGCAGAAAGTCATTTTCGTAATTCATGCGCTCATTTGCGATTTCC 564
QY 217 ThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeu 236
Db 565 AATTTTCGAAGTGTGTTTCTACCAACATATGCAACAGCTGCAATACACATTTATTGCTA 624
QY 237 LeuGlnGlnGlyAlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIle 256
Db 625 TTAAGATGCTCAAGTTTGGGAAGAAATGGGATATCT----- 666
QY 257 GluProAsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLys 276
Db 667 -----TCAGAAAGATGTTGCTGAATTTATCATAGACAAATTAACCTTACA---CAACAA 717
QY 277 TyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro 296
Db 718 TACACTGACCATTTGTGTTAATTTGTTAATTTGTTAATTTGTTAATTTGTTAATTTGTTA 777
QY 297 AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAsp 316
Db 778 TATGATGATGGGTCAAATTTAAACCGTTTTCGACAGAAATGACTTTAACTGTATTAGAT 837
QY 317 ThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIle 336
Db 838 CTAATTTGTACTTTTCCCATTTTATGATTTGCGTTATCTCA----- 879
QY 337 GluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPhe 356
Db 880 -----AAAGGGGTAAACAGAACTAAACAGAGACATTTTACGGATCCAATTTTCT 933
QY 357 AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThr 376
Db 934 CTTAGGACACCA---CTTGGGTACGGACCAACTTTTGTAGTATAGAAAACTCTATTTCGA 990
QY 377 ArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr----- 393
Db 991 AAACCT-----CATTTATTGATTATTATCAGGGGATGAAATTCATACGCGCTTCAA 1044
QY 394 -----GluAsnThrAsnPhe-----GlyAsnArgLeuValGlyIle 405

Db 1045 CCTGGTTACTTTGGAAAGATCTTTCATATTATTGGTCTGGTAATAT-----GTA 1095
QY 406 SerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArg 425
Db 1096 GAAACTAGACCTAGTATAGGATCTAGTAAGACAACTTACTTCCCATTTTATGGAGATAAA 1155
QY 426 ThrGlySerProThrThrLysThrIleArgProPheGluSer-----TyrLysVal 442
Db 1156 TCTACTGAACCTGTACAAAGCTA-----AGCTTTGATGGACAAAAGATTATCGAACT 1209
QY 443 SerIleValThrArgGlnSerProPro-----ValSerPro 455
Db 1210 ATACTAATACAGACCTAGCGCTTGGCCGAATGTAAGGTATATTTAGGTGTTAGGAAA 1269
QY 456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn 475
Db 1270 GTTGAT-----TTTAGTCAATATGATGATCAAAAAAATGAAACTAGTACACAA 1317
QY 476 ThrLeuLysTyrSerAlaGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGln 495
Db 1318 ACATATGATTTCAAAAAGAAACAATGGCCATGTAAGTGACACAGGATTTCTATTGACCAATTA 1377
QY 496 PheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsn 515
Db 1378 CCGCCAGAAACAACAGAT-----GAACCATCTGAAAAA 1410
QY 516 AsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeu 535
Db 1411 GCATATAGTCATCAGCTTAATACGCGAATGTTCTTA-----ATG 1452
QY 536 GlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyr 555
Db 1453 CAGGACCGTCGTGGAACAATTCATTTTCTTGGACACATAGAAAGTGTAGACTTTT 1512
QY 556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp 575
Db 1513 AATCAATGATGCTGAAAGATTTACTCACTTCAGTAGTGAAGCATATGCCTTGTCT 1572
QY 576 ThrAsnSerLysIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln 595
Db 1573 TCAGGTGCTCCATATTGAAGTCCAGGATTCACAGGAGGAAATTTACTATTCTCTAAA 1632
QY 596 -----SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGln 611
Db 1633 GAATCTAGTAATTCATTTGCTAAATTTAAAGTTTACATTAATTCAGCAGCTTGTACAA 1692
QY 612 SerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIle 631
Db 1693 CGATATCGTGTAAAGATACGCTATGCTCTTCTACC-----ACTAATCTTA 1734
QY 632 SerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsnAsnThrPheSer 651
Db 1735 CGACTTTTGTG-----CAAAATTCAAACATGATTTTCTT 1770
QY 652 GlyThrAsnTyrAsn-----AsnLeuGlnTyrGlyAspPheGly 664
Db 1771 GTCATCTACATTAATAAACTATGAATAAGATGATGATTTTAAACATATCAACATTTGAT 1830
QY 665 TyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhePhePhe 684
Db 1831 CTCGCAACTACTAATTTCTAATATGGGTCTCGGGTGATAAGAAATGAACCTATAATAGGA 1890
QY 685 AsnArgAlaAspValSerAsnSerIleLeuIleAspLysIleGluPheIleProfile 704
Db 1891 CGAATCTTTCGTTTCTTAATGAAAAAATCTATATAGATAGATAGAAATTTATCCAGTA 1950
RESULT 12
US-08-953-775B-37
; Sequence 37, Application US/08993775B
; Patent No. 6077824
; GENERAL INFORMATION:
; APPLICANT: English, Leigh H.

298	Db	TGCGCAAGT-----GATGCTGACCCATGGAGGCGTTTATGGCACAAGTTGAAGTA	348
137	Qy	PheValAspThrProLeuThrGluSerIleLeysGlnLeuLysLeuGlnThrLeuGluGly	156
349	Db	CTGATAGATAAGAAATAGAGGAGTATGCTTAAAAGTAAAGCTCTTGACAGAGTTACAGGCT	408
157	Qy	PheArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArg	176
409	Db	CTTCAAAATAATTTTCGAAGATTATGTTAATGTCGGTTAAATTCCTGGAGAATA-----	459
177	Qy	LeuGlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLys	196
460	Db	-----ACACCTTTTAAGTTTCGGAATAAAAGAACAGATCGAATAAGG	504
197	Qy	IleArgPheGluAsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGlu	216
505	Db	GAACCTTTTCTCAAGCAGAAAGTCATTTTCGTAATTCATGCGCTCATTTGCAGTTTCC	564
217	Qy	ThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeu	236
565	Db	AAATTTCGAAGTCTGTTTCTACCAACATATGCACAAGCTGCAAAATACACATTATTGCTA	624
237	Qy	LeuGlnGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIle	256
625	Db	TTAAAAGATGCTCAAGTTTTTGGAGAGAATGGGGATTTCT-----	666
257	Qy	GluProAsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLys	276
667	Db	-----TCAGAAGATGTTGCTGAATTTTATCATAGACAATTAAACCTTACA---CAACAA	717
277	Qy	TyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro	296
718	Db	TACACTGACCATGTTGTTAATTCGTATATGTTGGATTAAATGGTTTAAGAGGTTCAACT	777
297	Qy	AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAsp	316
778	Db	TATGATGCATGGGTCAAAATTTAAACCGTTTTCGCAGAGAAATGACTTTTAACTGATTAGAT	837
317	Qy	ThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIle	336
838	Db	CTAATTGTACTTTTCCCATTTTANGATTATTCGGTTTACTCA-----	879
337	Qy	GluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPhe	356
880	Db	-----AAGGGGTAAACACAGAACTTACAAAGACACATTTTACGGATCCAAATTTTCT	933
357	Qy	AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThr	376
934	Db	CTTAGGACACCA---CTTGGCTACGGACCACTTTTTTTGGATATAGAAAACTCTATTTCGA	990
377	Qy	ArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr-----	393
991	Db	AAACCT-----CATTTATTATTGATTATTTACAGCGGAATGAAATTCATACGCGTTCCAA	1044
394	Qy	-----GluAsnThrAsnPhe-----GlyAsnArgLeuValGlyIle	405
1045	Db	CCTGGTTACTTTGGGAAAGATTTCTTCAATATATGGTCTGTAATTAT-----GTA	1095
406	Qy	SerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArg	425
1096	Db	GAAACTAGACCTAGTATAGGATCTAGTAAGACAATTAATCTCCCATTTTATCGAGATAAA	1155
426	Qy	ThrGlySerProThrThrLysThrIleArgProPheGluSer-----TyrLysVal	442
1156	Db	TCCTACTGAACCTGTACAAAAGCTA-----AGCTTTGATCGCAAAAAGTTTATCGAAT	1209
443	Qy	SerIleValThrAspArgGlnSerProPro-----ValSerPro	455
1210	Db	ATAGCTAATATACACAGCTAGCGGCTTGCCCAAGTAAGGTATATTTAGGTGTTACGAAA	1269
456	Qy	IleGlnProHisPheIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn	475
1270	Db	TTTGAT-----TTTAGTCAATATGATGATCAAAAATGAACTAGTACACAA	1317

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/427,769
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/993,722
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Kitchell, Barbara S.
 REGISTRATION NUMBER: 33,928
 REFERENCE/DOCKET NUMBER: MECO:149
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 512/418-3106
 TELEFAX: 512/474-7577
 INFORMATION FOR SEQ ID NO: 37:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1959 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1956
 US-09-427-769-37

Alignment Scores:
 Pred. No.: 2,58e-60 Length: 1959
 Score: 663.00 Matches: 208
 Percent Similarity: 43.51% Conservative: 114
 Best Local Similarity: 28.11% Mismatches: 292
 Query Match: 17.14% Indels: 126
 DB: 3 Gaps: 27

US-10-783-417-2 (1-735) x US-09-427-769-37 (1-1959)

```

QY 1 MetAsnGlnAsnAspAsnAsnGluTyrGluIleAspSerHisThrSerProTyr 20
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  1 ATGAATCCAAACAT---CGAAGTGAACATGAT-----ACGATAAAGTT 42
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 21 PheProAsnArg-----AsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsn 38
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  43 ACACCTAACAGTGAATGTCACAACTAACCAATAATCAATATCCTTTAGCTGACAAATCCAAT 102
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 39 GlnProLeuGlnAsnThrAsnTyrLysGluTrpLeuAsnMetCysGln---GlyAsnThr 57
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  103 TCAACACTAGAAGATTAAATTTATAAAGATTTTTAAAGATGACTGAAGACAGTTCCTACG 162
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 58 GlnTyrGlyAspAsnPheGluThrPheAlaSerAlaAspThrIleAla---AlaValSer 76
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  163 GAAGTCTAGAACAC-----TCTACAGTAAAGATGACAGTGGG 201
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 77 AlaGlyThrIleValSerGlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSer 96
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  202 ACAGGAATTCGTGTAGGCGAGATTAGGTGTGTAGGA----- 243
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 97 GlyProIleGlyIleLeuAlaIleIleSerPheGlyThrLeuIleThrValPhe 116
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  244 -----GTTCCATTGCTGGGCACTCACTTCATTTTATCAATCATTTCTTAACACTATA 297
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 117 TrpProAlaGlyGluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIle 136
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  298 TGGCCCAAGT-----GATGCTGACCCCATGGAAGCTTTTATGGCACAAGTTGAAGTA 348
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 137 PheValAspThrProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGly 156
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  349 CTGATAGATAAGAAAATAGAGGATGATGCTAAAGTAAAGTAAAGCTTTGACAGAGTTACAGGGT 408
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 157 PheArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArg 176
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 409 CTTCAAAATAATTTTCGAAGATTATGTTAATGCGTTAAATCTCTGGAGAAA----- 459
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 177 LeuGlnAlaProGlyLeuProProSerAlaLeuGlnGlnAlaLeuThrLeuLys 196
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 460 -----ACACCTTTAAGTTTGCAGAGTAAAGAAAGCAAGCAAGTCAATAAGG 504
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 197 IleArgPheGluAsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGlu 216
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 505 GAACCTTTTCTCAAGCAGAAAGTCAATTTTCGTAATTCATCCGTCATTTGCGAGTTTC 564
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 217 ThrTyrLysThrLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeu 236
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 565 AAATTCGAAGTCTGTTTCTACCAACATATGCACAAAGCTGCAAAATACACATTTATTGCTA 624
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 237 LeuGlnGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIle 256
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 625 TTAAGAGATGCTCAAGTCTTTTGGAGAGAAATGGGATATTCT- 666
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 257 GluProAsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLys 276
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 667 -----TCAGAGAGATGTTGCTGAATTTTATCATAGACAAATTAATAACTTACA---CAACAA 717
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 277 TyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro 296
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 718 TACACTGACCAATTGTTAATTTGGTATTAATGTTGAATTAATGGTTTAAAGAGTTCAACT 777
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 297 AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAsp 316
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 778 TATGATGCATGGGTCAAATTTAACCGTTTTGCAGAGAAATGACTTTAACATGTATTAGAT 837
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 317 ThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIle 336
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 838 CTAATTTGACTTTTCCCATTTTATGATATTCGGTTATACTCA----- 879
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 337 GluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPhe 356
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 880 -----AAAGGGGTAAACAGAACTAACAGAGACATTTTACGGATCCCAATTTTCT 933
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 357 AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThr 376
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 934 CTTAGGACACCA---CTTGCCTACGGACCACTTTTGTAGTATAGAAAACCTCTATTCCA 990
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 377 ArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr----- 393
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 991 AAACCT-----CATTTATTGATTATTACAGGGGATTGAATTGATCATACGGCTCTTCAA 1044
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 394 -----GluAsnThrAsnPhe-----GlyAsnArgLeuValGlyIle 405
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1045 CTGTTTACTTTGGGAAAGATTCTTCAATTTATTGGTCTGGTAAATTAT-----GTA 1095
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 406 SerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArg 425
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1096 GAACCTAGACTAGTATAGGATCTAGTAAGACAATTACTTCCCAATTTTATGGAGATAAA 1155
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 426 ThrGlySerProThrThrLysThrIleArgProPheGluSer-----TyrLysVal 442
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1156 TCTACTGAACCTGTACAAAAGCTA-----AGCTTGTAGTGCACAAAAGTTTATCGAACT 1209
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 443 SerIleValThrAspArgGlnSerProPro-----ValSerPro 455
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1210 ATAGCTAATACAGACGTAGCGCTTGGCCGAATGGTAAGGTATATTAGTGTGTACGAAA 1269
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn 475
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1270 GTTGAT-----TTTAGTCAATATGATGATCAAAAATAAGAACTAGTACACAA 1317
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 476 ThrLeuLysTyrSerAlaGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGln 495
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1318 ACATATGATCAAAAAGAAACAAATGGCCATGTAAGTGCACAGGATTTCTATTACCAATTA 1377
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 496 PheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsn 515
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```
Db 1378 CCGCCAGAACACAGAT-----GAACACCTTGAAAAA 1410
QY 516 AsnTyrSerHisIleLeuSerHisPheThrTyrSerTyrValIleGlyLeu 535
Db 1411 GCATATAGTCATCAGCTTAATACGGGAATGTTCTTA-----ATG 1452
QY 536 GlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrHisSerSerValAspArgTyr 555
Db 1453 CAGGACCGTGTGGAAACAATTCATTTTACCTGGACACATAGAGTGAGACTTTTT 1512
QY 556 AsnAlaIleSerAspIleIleThrMetIleProAlaIleLeuGlyGlyAsnAsnLeuAsp 575
Db 1513 AATACAAATTCAGTCTGAAAGATATCTCAACTCCAGTAGTAGAAGCATATGCTTGTCT 1572
QY 576 ThrAsnSerIysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln 595
Db 1573 TCAGGTGCTTCCATTTATGAAGTCCAGGATTCACAGGAGAAATTTACTATTCTTAA 1632
QY 596 -----SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGln 611
Db 1633 GAATCTAGTAATTCATTTGCTAAATTTAAAGTTTACATTAATTCAGCAGCCTTGTTACAA 1692
QY 612 SerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIle 631
Db 1693 CGATATCGTGAAGATACGCTATGCTTCTACC-----ACTAACTTA 1734
QY 632 SerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsnAsnThrPheSer 651
Db 1735 CGACTTTTGTG-----CAAAATTCAAACAATGATTTCCTT 1770
QY 652 GlyThrAsnTyrAsn-----AsnLeuGlnTyrGlyAspPheGly 664
Db 1771 GTCATCTACATTAATAAACTATGAATAAAGATGATGATTTAAACATATCAACATTTGAT 1830
QY 665 TyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePhe 684
Db 1831 CTGCGAACTACTAATTTCTAATATGGGGTCTCGGGTGATAAGAATGAACATTATAATAGGA 1890
QY 685 AsnArgAlaAspValSerAsnSerIleLeuIleIleAspIysIleGluPheIleProIle 704
Db 1891 GCAGAACTCTTCGTTTCTTAATGAANAATCTATATAGATAAGATAGAAATTTATCCCGATGA 1950

RESULT 15
US-07-973-320-3
; Sequence 3, Application US/07973320
; Patent No. 5286486
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel M.
; APPLICANT: Fu, Jenny M.
; TITLE OF INVENTION: No. 5286486el Bacillus thuringiensis Gene
; TITLE OF INVENTION: Encoding a Coleopteran-Active Toxin
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/973,320
; FILING DATE: 19921106
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/788,638
; FILING DATE: 6-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
```

```
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA68.C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3414 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: kumamotoensis
; INDIVIDUAL ISOLATE: HD867
; IMMEDIATE SOURCE:
; LIBRARY: Lamdagem (TM)-11 library of J.M. Fu
; CLONE: 867
; US-07-973-320-3

Alignment Scores:
Pred. No.: 1.18e-59 Length: 3414
Score: 660.50 Matches: 223
Percent Similarity: 43.06% Conservative: 112
Best Local Similarity: 28.66% Mismatches: 286
Query Match: 17.07% Indels: 157
DB: 2 Gaps: 37

US-10-783-417-2 (1-735) x US-07-973-320-3 (1-3414)
QY 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 1 ATGAATTTAATAATTTAGTGGA---TATGAA-----GATAGT----- 36
QY 21 PheProAsnArgAsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGln 39
Db 37 -----AATAGAACATTAATAATCTCTCAATTTATCTCTCAAAAGACATTATCACCA 90
QY 40 ProLeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyr 59
Db 91 TCATTAAGAATATGAATCACTACAGGATTTTATCTATACTAG----- 135
QY 60 GlyAspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThr 79
Db 136 AGGGAACAGCCTGAAGCAGCTCGTAGTGGTAATACA-----GCTATTAACTAGTAGTT 189
QY 80 IleValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIle 99
Db 190 AGTGTACAGGGGCTACACTAAGTCGCTAGT-----GTCCCAGGT----- 231
QY 100 GlyIleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThr---ValPheTrpPro 118
Db 232 -----GCAAGTTTTATCACTTAACCTTTTACCTGAAATTTACAGGCTTTTATGGCCA 282
QY 119 AlaGlyGluGlnAspIysThrValTyrThrGlnPheIleIleLysMetGlyGluIlePheVal 138
Db 283 -----CACGATAAAATAATTTGGGATGAATTTATGACAGAGTAGAACAACCTTATT 333
QY 139 AspThrProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArg 158
Db 334 GAACAAAAATAGACAATATGCAAGGAATAAAGACACTTCGCAAGTATAGAGGATTAGGA 393
QY 159 GlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGln 178
Db 394 AATAACTTAACGATATATCAACAGGCACCTTGAAGATTGG-----CTGAAC 438
QY 179 AlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIle--- 197
Db 439 AATCCTGATGATCCA-----GCAACTATAACACGAGTAGATAGAT 477
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 11, 2005, 17:40:06 ; Search time 1443 Seconds
(without alignments)

4212.053 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869

Sequence: 1 MQQNDNVEYIDSHSTPY.....KLETIQKINTFTNHTKTL 735

Scoring table: BLOSUM62

Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10783417/runat_01122005_141955_24239/app_query.fasta_1.903
-DB=Published Applications NA.Main -OPMT=fastap -SUFFIX=rnpbm -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blsum62
-TRANS-human40.cdi -LIST=45 -DLOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10783417 @CGN 1 1 1026 @runat_01122005_141955_24239
-NCPU=6 -ICPU=3 -NO MAP -LARGESUBSTR -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXI=7

Database :

Published Applications NA.Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3869	100.0	2208	8 US-10-783-417-1	Sequence 1, Appli
2	3314	85.7	2235	8 US-10-782-570-1	Sequence 1, Appli
3	3104.5	80.2	2085	8 US-10-782-570-3	Sequence 3, Appli
4	849	21.9	4896	3 US-09-756-526A-3	Sequence 3, Appli
5	849	21.9	4896	6 US-10-345-020-3	Sequence 3, Appli
6	849	21.9	4896	6 US-10-342-821-3	Sequence 3, Appli
7	759	19.6	3684	9 US-10-929-754-2	Sequence 2, Appli
8	705	18.2	4391	9 US-10-929-754-4	Sequence 4, Appli

9	703	18.2	3633	5	US-10-032-717-3	Sequence 3, Appli
10	703	18.2	3633	6	US-10-414-637-3	Sequence 3, Appli
11	703	18.2	3633	7	US-10-606-320-3	Sequence 3, Appli
12	703	18.2	3633	8	US-10-746-914-3	Sequence 3, Appli
13	703	18.2	6613	5	US-10-032-717-28	Sequence 28, Appli
14	703	18.2	6613	6	US-10-414-637-28	Sequence 28, Appli
15	703	18.2	6613	7	US-10-606-320-18	Sequence 18, Appli
16	703	18.2	6613	8	US-10-746-914-18	Sequence 18, Appli
17	689.5	17.8	4188	5	US-10-120-544A-5	Sequence 5, Appli
18	689.5	17.8	4188	10	US-11-091-654-5	Sequence 5, Appli
19	678.5	17.5	2003	5	US-10-032-717-7	Sequence 7, Appli
20	678.5	17.5	2003	6	US-10-414-637-7	Sequence 7, Appli
21	678.5	17.5	2003	5	US-10-032-717-17	Sequence 17, Appli
22	678.5	17.5	2022	5	US-10-414-637-17	Sequence 17, Appli
23	678.5	17.5	2022	6	US-10-606-320-13	Sequence 13, Appli
24	678.5	17.5	2022	8	US-10-746-914-13	Sequence 13, Appli
25	675	17.4	3621	5	US-10-032-717-1	Sequence 1, Appli
26	675	17.4	3621	6	US-10-414-637-1	Sequence 1, Appli
27	675	17.4	3621	7	US-10-606-320-1	Sequence 1, Appli
28	675	17.4	3621	8	US-10-746-914-1	Sequence 1, Appli
29	675	17.4	3621	10	US-11-021-115-5	Sequence 5, Appli
30	675	17.4	4874	5	US-10-032-717-27	Sequence 27, Appli
31	675	17.4	4874	6	US-10-414-637-27	Sequence 27, Appli
32	675	17.4	4874	7	US-10-606-320-17	Sequence 17, Appli
33	675	17.4	4874	8	US-10-746-914-17	Sequence 17, Appli
34	669	17.3	4359	5	US-10-120-544A-3	Sequence 3, Appli
35	669	17.3	4359	10	US-11-091-654-3	Sequence 3, Appli
36	668	17.3	2088	10	US-11-018-615-5	Sequence 5, Appli
37	663	17.1	1959	7	US-10-614-076-37	Sequence 37, Appli
38	663	17.1	2358	10	US-11-018-615-28	Sequence 28, Appli
39	663	17.1	3510	10	US-11-018-615-24	Sequence 24, Appli
40	660	17.1	1959	5	US-10-232-665-3	Sequence 3, Appli
41	658	17.0	1959	7	US-10-614-076-27	Sequence 27, Appli
42	658	17.0	1959	7	US-10-614-076-35	Sequence 35, Appli
43	657.5	17.0	2025	7	US-10-606-320-73	Sequence 73, Appli
44	657.5	17.0	2025	7	US-10-606-320-79	Sequence 79, Appli
45	657.5	17.0	2025	8	US-10-746-914-73	Sequence 73, Appli

ALIGNMENTS

RESULT 1
US-10-783-417-1
; Sequence 1, Application US/10783417
; Publication No. US20040216186A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Kozziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-006, A Delta-Endotoxin Gene and
; FILE REFERENCE: 045600/274146
; CURRENT APPLICATION NUMBER: US/10/783,417
; PRIOR FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: 60/448,806
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2208
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2208)
US-10-783-417-1
Alignment Scores:
Pred. No.: 0 Length: 2208
Score: 3869.00 Matches: 735
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-10-783-417-2 (1-735) x US-10-783-417-1 (1-2208)

```
QY 1 MetAsnGlnAsnAspAsnAsnGluTyrGluIleAspSerHisThrSerProTyr 20
DB 1 ATGAATCAAAATACCATATTAACGAATATGAATATTAATGATTCGCATACCTCACCTTAT 60
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
DB 61 TTTCCGAAACAGAAACAGTATGATTTCTAGATACCTTACAAATATATCCAAATCAACCA 120
QY 41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
DB 121 TTTCAAAACACAAATTTACAAAGAGTGGCTCAATATGTGTCAAGGGAATACACAATATGGT 180
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIle 80
DB 181 GATAATTTTCGAGACATTTGCTAGTGTGTATACAAATTTGCTGAGTTAGTGCAAGTACTATT 240
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
DB 241 GTATCCGGTACTCTGTGTAGCCGTATAGGTGGCTCACTTCTATATCCGACCGATAGGA 300
QY 101 IleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGly 120
DB 301 ATAATAGGTGCTATATATATCTTTTGGTACCTTAATCACTGCTCTTTGGCCCGGGGA 360
QY 121 GluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
DB 361 GAACAAGACAAACAGTATGGACACAAATTTATTAATAATGGAGAAATTTTGTGTGATACA 420
QY 141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
DB 421 CCGTTAAACAGAAAGCATAAACAGCTTAAAGTTTACAAACTTTTAGAAGGATTTAGACAAATA 480
QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaPro 180
DB 481 TTACAAGCTATAATACAGATTAATGATATGATGGAGAAATTTAAAGACTTACAAAGTCTCT 540
QY 181 GlyLeuProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGlu 200
DB 541 GGATTTACCAACCATCATCAGCATTTACAAAGCTGCTTGACTCTTAAATACGATTTAG 600
QY 201 AsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThr 220
DB 601 AATGTTCACAATGATTTTATTCGAGAAATACCTGGTTTCCAACTTGAAACTTATAAAGC 660
QY 221 LeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGly 240
DB 661 CTATTACTACCTATTATGCGCAAGCTGCTAAATTTTCAATTTTATTAACCAAGGT 720
QY 241 AlaGluLeuAlaAspGluTyrPheAlaAspIleHisProSerGlnIleGluProAsnAla 260
DB 721 GCTGAATGCGCTGATGAATGGAATGAGATATACATCTTCAAAATTTGAACCTAATGCT 780
QY 261 GlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyr 280
DB 781 GGAAATCAGATGACTATTATTAACCTTTTAAAGAAAAATATATACCTAAATATAGTAACAT 840
QY 281 CysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTyr 300
DB 841 TGTGCAAAATACCTTATAGACAGGACTTAAAGAAATCTTTAGAGACCAATATGAATATGG 900
QY 301 SerIlePheAsnAspTyrArgTyrMetThrIleThrValLeuAspThrIleSerGln 320
DB 901 AGTATATTTAATGACTATCGAAGATATATGACCATTTACTGTATTAATACCATCTCTCAA 960
QY 321 PheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyIleGluValLysGly 340
DB 961 TTTTCTTTTATATGATATAAAGATATAGAGATTTCAATAGGAGGAATAGAAAGGC 1020
```

[illegible]

RESULT 2

```

US-10-782-570-1
; Sequence 1, Application US/10782570
; Publication No. US20040210965A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-007, A Delta-Endotoxin Gene and
; TITLE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/274144
; CURRENT APPLICATION NUMBER: US/10/782,570
; CURRENT FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: 60/448,812
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2235
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2235)
US-10-782-570-1

```

Qy	141	ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle	160
Db	418	CCGTTAA CAGAAGCATAAACACGCTAAAGTTACAAACTTTTAGAAGGATTTAGACAAATA	477
Qy	161	LeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaPro	180
Db	478	TTACAAAGCTATTAAATACAGCATTAGATGTGGAGAAAAATTAAAGAACGTACAAAGCTCCT	537
Qy	181	GlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGlu	200
Db	538	GGATTACCACCATCATCAGCATTACAAACAGCTGCCTTTCATCTTTAAATATACGATTTGAG	597
Qy	201	AsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThr	220
Db	598	AATGTTTCCAAATGATTTTATTCGAGAAATACCTGGTTTCCAACTTGAAACTTATAAAACG	657
Qy	221	LeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGly	240
Db	658	CTATTACTACTATTATGCGCAGCTGCTAAATTTTCATTTAAATTTATTACAAACAGGT	717
Qy	241	AlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAla	260
Db	718	GCTGAATTGCTGATGATGAAATGCAGATATACATCTTCACAAATTGAACCTAATGCT	777
Qy	261	GlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyr	280
Db	778	GGAAACATCAGATGACTATTATAAACTTTTAAAGAAAAATATACCTAAATATAGTAACTAT	837
Qy	281	CysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTyr	300
Db	838	TGTGCAAAATACCTATTAGAGAAGACATAAATAAACTTCGAAACGAACTAATATGAGATGG	897
Qy	301	SerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGln	320
Db	898	AGTATATTTAATGATATCGAGATATAGCATTTACTGTATTAGATACATATCGCTCAA	957
Qy	321	PheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluValLysGly	340
Db	958	TTTTTCTTTTATGATATAAGAGATACAAAGATTCATAGGA-----AGAATAGTGGC	1011
Qy	341	IleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuPro	360
Db	1012	ATTAAAACTGAACTTACACAGAGAAATTTACAACTGAAATAAATTTTGACCGCTTACT	1071
Qy	361	GlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPhe	380
Db	1072	TACCTTGAATTTCAACCCCAATCTCGCTATAATGGAATATAAATTTAACACGTTCAGGCGCT	1131
Qy	381	LysLeuPheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsn	400
Db	1132	AGATTATTTTCATTTTGTAGTAGAATTATATTTTATACAAAAAATGAAACGTACGGGAAT	1191
Qy	401	ArgLeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr	420
Db	1192	CGTTTAGTTGGTATTGCGAATCGTAATAGATCTACTTATGCTACGACAGGAACTGAAATT	1251
Qy	421	LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr	440
Db	1252	ATATATGGAGAAAGACAGGTCACCCCAACAAAAACTTTTATACCAATTTGAATTCCTAT	1311
Qy	441	LysValSerIleValThrAspArgGlnSerProProValSerProIle---GlnProHis	459
Db	1312	AAAGTTTCATTTGTAAGTAGAAGTAAGTAACCTCTACTTCCCTTTTCTTAACATATAC	1371
Qy	460	PheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnThrLeuLysTyr	479
Db	1372	TTTTACAATTAATCAAATTTGAACCTTTATTTAAATAATTCACCTAGTAATAAATAACATAT	1431
Qy	480	SerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArgLys	499
Db	1432	TCAGCTGGGGGAAATTTATCTAATGATAAAAAAACCACTGATTTTCAATTTCTGTAAAA	1491

```
QY 500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHis 519
Db 1492 AAAGACTGTAACCAATATTATTAACCAATGTTTACCAAGCTAATATAGTTATAGTCAT 1551

QY 520 IleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIle 539
Db 1552 ATTTATCCAGTTTCTTTTATTATTATCTCTATAAAATTCGATAGCGCTTAATATA 1611

QY 540 LeuAspThrGlyValLeuGlyThrHisSerSerValAspArgTyrAsnAlaIleSer 559
Db 1612 TTATATACAGGTGCATTAGGATGCACACAGTAGTGTTAATAGAAATAATGCAATATCA 1671

QY 560 AspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLys 579
Db 1672 GATAAAATAATTAACAATGATCCAGCAATCAAAAGGTAACAGACTCTTGATACAAACTCTAAG 1731

QY 580 ValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArg 599
Db 1732 GTAATTGAAGGACCTGTCATACAGAGGAACTTGTTTATTATTAACAAAGTCAAGGCGT 1791

QY 600 LeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyr 619
Db 1792 TTAGAGATTACATGTAGAACTCTTAATCTTACACAATCTTATTACATTAGACTTCGATAC 1851

QY 620 AlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerIleuThrIleProGlyValIle 639
Db 1852 GCTACAAATGGTGCTGGAAATCTCTTCTTAATATATCTCTTACAAATACAGGAGTAATA 1911

QY 640 GlyIleProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGln 659
Db 1912 GGAATACCACTCAACGACTCAACAACTTTTCTGGTACAAATTAATAATTTACAA 1971

QY 660 TyrGlyAspPheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsn 679
Db 1972 TACGGAGATTTGGGTATTTCCAAATTCAGTACAGTAACTTACCTTTAAATCGAATC 2031

QY 680 IleProPheIlePheAsnArgAlaAspValSerAsnSerIleLeuIleAspLysIle 699
Db 2032 ATACCAATTTATATTTAATCGTCAGATGTATCAAAATTCAAATTTAATCATTGATAAAAT 2091

QY 700 GluPheIleProIleThrSerSerMetHisGlnAsnArgGlnLysGlnLysLeuGluThr 719
Db 2092 GAATTTATACCAATTAATCTCTCTGTCAGCAATGTATCAAAATTCAAATTTAATCATTGATAAAAT 2151

QY 720 IleGlnThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db 2152 ATCCAAACAAATAATAATACATTTTTCACAAATCATACAAA 2193
```

RESULT 3

```
US-10-782-570-3
; Sequence 3, Application US/10782570
; Publication No. US20040210965A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargies, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-007, A Delta-Endotoxin Gene and
; TITLE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/274144
; CURRENT APPLICATION NUMBER: US/10782,570
; CURRENT FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: 60/448,812
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2085
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: CDS
```

```
; LOCATION: (1)...(2085)
US-10-782-570-3
```

Alignment Scores:

```
Pred. No.: 0 Length: 2085
Score: 3104.50 Matches: 598
Percent Similarity: 91.95% Conservative: 30
Best Local Similarity: 87.55% Mismatches: 52
Query Match: 80.24% Indels: 3
DB: 8 Gaps: 2
```

US-10-783-417-2 (1-735) x US-10-782-570-3 (1-2085)

```
QY 52 MetCysGlnGlyAsnThrGlnTyrGlyAspAsnPheGluThrPheAlaSerAlaAspThr 71
Db 1 ATGTGTCAAGGGAATACACAATATGGTGATAATTTTCGAGACATTGCTAGTGTGATACA 60

QY 72 IleAlaAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAlaGlyIleGlyGly 91
Db 61 ATTGCTGCAGTTAGTGCAGGTACTATTGTATCCGGTACTCTGTTAGCCGGTATAGTGGG 120

QY 92 LeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIleIleSerPheGlyThr 111
Db 121 CTCACCTTCTATATCCGGACCGATAGGAATATAGGTGCTATATAATATCTTTTGGTACC 180

QY 112 LeuIleThrValPheTrpProAlaGlyGluGlnAspLysThrValTrpThrGlnPheIle 131
Db 181 CTAATCACTGCTTTTGGCCCGGGGAGAACAGACAAACAGTATGGACACAAATTTATT 240

QY 132 LysMetGlyGluIlePheValAspThrProLeuThrGluSerIleLysGlnLeuLysLeu 151
Db 241 AAAATGGGAGAAATTTTGTGTATACACCGTTTAAACAGAAAGCATAAACAGTAAAGTTA 300

QY 152 GlnThrLeuGluGlyPheArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTrp 171
Db 301 CAAACITTTAGAGGATTTAGACAAATATTAACAAAGCTTATATACAGCATTAGATGATTGG 360

QY 172 ArgLysLeuLysArgLeuGlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAla 191
Db 361 AGAAATTAATAAAGACTTACAAAGCTCTCGGATTACCACCATCATCAGCATTTACAACAGCT 420

QY 192 AlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPheIleArgGluIlePro 211
Db 421 GCCTTGACTCTTAAATATACGATTTTGAGATGTTTGCATGATGATTTTATTCGAGAAATACCT 480

QY 212 GlyPheGlnLeuGluThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAsn 231
Db 481 GGTTCCTCACTTGAAACTTATAAAGCGCTATTACTACTATTTATGCGCAAGCTGCTAAT 540

QY 232 PheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIle 251
Db 541 TTTTCATTTAAATTTATTAACAAGGCTGCTGAATGGCTGATGATGATGATGATGATGATGAT 600

QY 252 HisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTyrTyrLysLeuLys 271
Db 601 CATCCTTCAAAATTTGAACCTTAATGCTGGAACATCAGATGATGATGATGATGATGATGATGAT 660

QY 272 GluAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsn 291
Db 661 GAAATATATACCTAAATATAGTAATCTATTGCAAAATACCTATAGAGAGGACTAAATATAA 720

QY 292 LeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThr 311
Db 721 CTTTGAACCAACCACTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780

QY 312 IleThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAsp 331
Db 781 ATTACTGATTAGATACTATCGCTCAATTTCTTTTATGATATAAAGAGATACAAAGAT 840

QY 332 SerIleGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThr 351
Db 841 TCAATAGGA-----AGAATAGGTGGCATTTAAACCTGAACCTTACAGAGAAATTTATACA 894
```


1369 TGCTCAAAATTTATGAACACGAGAGACCTTTTAAATCAAAATTTCTACAGCTGTA 1428
Db
147 LysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGlnSerTyrAsnThr 166
Qy
1429 AAGAAATAGCATTAGCTCATCTAAATGGTTTAAAGATGATTAATACGTACTATAAGA 1488
Db
167 AlaLeuAspAspTrpArgLysLeuLysLeuArgLeuGlnAlaProGlyLeuProSerSer 186
Qy
1489 GCATTAATGATGGAAGAGA-----AATCCAAGTGCA 1521
Db
187 AlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPhe 206
Qy
1522 -----AATAGTCCAGATTGGTATCACAGATTGGAACGCTCATTTCAATTTT 1572
Db
207 IleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeuProIleTyr 226
Qy
1573 GTAAGCAATATGCAAACTCCAACCTCCCACTCCCACTGATGACATTTATTAAAGTTGCTAT 1632
Db
227 AlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGlnGlyAlaGluLeuAlaAspGlu 246
Qy
1633 ACAGAGCTGCAAAATTTACATTTGAATTTATTATCATCAAGGTGTACAAATTCGCGGATCAA 1692
Db
247 TrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTyr 266
Qy
1693 TGGAAATGCAGATCAACACCATTCACCAATGTTGAAGTCATCAGTACT-----TAT 1743
Db
267 TyrLysLeuLysGluAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyrArg 286
Qy
1744 TATGACGAGCTATTGGTATATATGAAAGTATATTAATTTATGACCAAGACATACCAT 1803
Db
287 ThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTyr 306
Qy
1804 AAGAGATTGAATCACCTTAAAGAAATCAGAAAAATCACATGGGATGCTTATAACACATAT 1863
Db
307 ArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIle 326
Qy
1864 CGTCGAGAAATGACCTTAATGTATTGGATCTTGTGCAACTTTTCCCTTTTATGATATA 1923
Db
327 LysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThr 346
Qy
1924 CGTCGTTTT-----CCAGAGAGTAGTAGAAGTAGAATTAACA 1959
Db
347 ArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeuArgValGlnPro 366
Qy
1960 AGAGAGTTTATACAAGTTTAGATCATTTAAACAGCACCA----- 2001
Db
367 AsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeu 386
Qy
2002 -----CGGCTATTACTTGGCTG 2019
Db
387 GluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIleSer 406
Qy
2020 TCAGATATTGAGTATACACGAGAGTGTGGAGAGCGCATTTATTATCAGGTATT--- 2076
Db
407 AsnArgAspAlaProThrTyrSerAsnThr-----IleThrGluThrLeuTyrGly 423
Qy
2077 ---CGAGAGCTAAATATATATATCTGTTAACTATTTTACGATGAAATAATTTATGGT 2133
Db
424 GluArgThrGly---SerProThrThrLysThrIleArgProPheGluSerTyrLysVal 442
Qy
2134 AATACAAATAGATTAAGTAGACGCTCATTTACATTTATTACAGCGCAA-----TATGACT 2190
Db
443 SerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIleIle 462
Qy
2191 CACTTAAGCAATAACCGCTCTTTTCAACAAATAGCTGGTATATAAATAGTTATACAGTTTA 2250
Db
463 AsnGlnIleGluLeuLysLeuAsnGlySerSerAsnAsnThrLeuLysTyrSerAlaGly 482
Qy
2251 ATTCAAAAATCGTATTACAACTTTTAAACACGATAAT-----GAATATCAAAAAAT 2304
Db
483 GlySerLeuSerAsnTyr-----GlnAsnThrThrPhePheGlnPheProArgLys 499
Qy
2305 TTTAATGTGAATAATCAAAATGAACCTCAAGAACTACA----- 2343
Db

Qy 500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHis 519
Db 2344 -----AATATCTCTAATGATTATGGTGTTCATCA-----AACAGCAAAAAATTCAAACAT 2391
Qy 520 IleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIle 539
Db 2392 AATTATCTCATTTTCCATTAATCATCCAC-----AAGTTAGAGTTT 2433
Qy 540 LeuAsp-----ThrGlyValLeuGlyTyrHisSerSerValAspArgTyr 555
Db 2434 GCTGAGTATTTTCACTCTATATTTGTCATTAGTTGGACACACATAGTTAACTCCCAA 2493
Qy 556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp 575
Db 2494 AATTATATATCAGAAAGTGTGAGTACACAAATCCCAATGGTAAAGCTTACCAAGTT--- 2550
Qy 576 ThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln 595
Db 2551 ACTAACAAATTCAGTTATTAAGAGGACAGGTTTACAGGTGGAGATTTAATAGAACTTCGT 2610
Qy 596 SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIle 615
Db 2611 GAT-----AAATGTTCTTAAATGTAAGCT---AGTCTTTAAAAAATACGCTATA 2661
Qy 616 ArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIle 635
Db 2662 AGTCTATTTTATGCTGCAAAATACGCAATAGCTGTATCAATAGACGTAGGTAGTCCGGA 2721
Qy 636 ProGlyValIleGlyIleProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyr 655
Db 2722 GCAGGAGTTCTA-----TTGCAACCTACTCTTTCTAGAAAAAGGAAC 2763
Qy 656 AsnAsn-----LeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSer 670
Db 2764 AATAATTTTACAATTCAGACCTTAATTAAGATTTTCAATATCATACACTTTTAGTT 2823
Qy 671 ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAsp----- 688
Db 2824 GATATTGAATTCACGAAAGTGAAGAAATTCATATTCATTTGAAGCGAGAGATGATTAT 2883
Qy 689 ValSerAsnSerIleLeuIleAspLysIleGluPheIleProIleThrSerSerMet 708
Db 2884 GAGGAGGAGTGTATCTTTTAAATGATAAATTAGAGTTTCAAACTATAGATGAAAAATAT 2943
Qy 709 HisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePhe 728
Db 2944 -----ACTAATGAATGAATTTAGAGAGGCAAGAAAGACGATGATGATTATTT 2994
Qy 729 ThrAsnHisThrLysThrLeu 735
Db 2995 ATAACGCAACAAACGCTTTG 3015

RESULT 5

US-10-345-020-3

; Sequence 3, Application US/10345020

; Publication No. US20030150018A1

; GENERAL INFORMATION:

; APPLICANT: Jana, Wojciechowska

; APPLICANT: Evgeny, Lewitin

; APPLICANT: Ludmila, Revina

; APPLICANT: Igor, Zalunin

; APPLICANT: Galina, Chestukhina

; TITLE OF INVENTION: NOVEL DELTA-ENDOTOXINS AND NUCLEIC ACID SEQUENCES CODING THEREFOR

; FILE REFERENCE: S-30913B

; CURRENT APPLICATION NUMBER: US/10/345,020

; CURRENT FILING DATE: 2003-01-15

; PRIOR APPLICATION NUMBER: US 60/175,158

; PRIOR FILING DATE: 2000-01-07

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 3

; LENGTH: 4896

```
/ TYPE: DNA
/ ORGANISM: Bacillus thuringiensis
/ FEATURE:
/ NAME/KEY: source
/ LOCATION: (1)..(4896)
/ OTHER INFORMATION: Bacillus thuringiensis supsp. finitimus strain VKPM B-1161 (-axon)
/ OTHER INFORMATION: :29337
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1129)..(4458)
/ OTHER INFORMATION: product: Cry28Aa1 delta-endotoxin
/ PUBLICATION INFORMATION:
/ AUTHORS: Wojciechowska, et al.
/ TITLE: Two novel delta-endotoxin gene families cry26 and cry28 from Bacillus
/ TITLE: thuringiensis ssp. finitimus
/ JOURNAL: FEBS Lett.
/ VOLUME: 453
/ ISSUE: 12
/ PAGES: 46-48
/ DATE: 1999-06-18
/ US-10-345-020-3

Alignment Scores:
Pred. No.: 1-54e-79 Length: 4896
Score: 849.00 Matches: 229
Percent Similarity: 48.76% Conservative: 106
Best Local Similarity: 33.33% Mismatches: 248
Query Match: 21.94% Indels: 104
DB: 6 Gaps: 25

US-10-783-417-2 (1-735) x US-10-345-020-3 (1-4896)
QY 68 SerAlaAspThrIleAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAla 87
Db 1210 AGCAGTGATACAGTCGCTGTAGTAAGCGCAGGAGTTCTAGTTGGGTACTACTAC 1269
QY 88 GlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIle 107
Db 1270 -----GCCTTTGTCATCATTTGTTTAATCCA-----GGTGTGGTACTTTATA 1308
QY 108 SerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGlnAsp---LysThrVal 126
Db 1309 TCATTTGGAAACCTTGGCTCCCGTCTTTTGGCTGATCCAGAGGAAGATCCAAAAAAT 1368
QY 127 TrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeuThrGluSerIle 146
Db 1369 TGCTCACATTTATGAACACGGAGAGACCTTTTAATCAACAATTTCTACAGCTGTA 1428
QY 147 LysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGlnSerTyrAsnThr 166
Db 1429 AAAGAAATAGCATTAGCTCATCTAAATGGTTTAAAGATGTATTAAAGTACTATGAAAGA 1488
QY 167 AlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeuProSerSer 186
Db 1489 GCATTAATGATTGGAAGAGA-----AATCCAAGTGCA 1521
QY 187 AlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPhe 206
Db 1522 -----AATACTGCCAATGGTATACAGAGATTTGAAACGCTCAITTCATTTT 1572
QY 207 IleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeuProIleTyr 226
Db 1573 GTAAGCAATATGCCACAACCTCAACTTCCACGATGATGACACATTTATTAAAGTTGCTAT 1632
QY 227 AlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGlu 246
Db 1633 ACAGAAGCTGCAAAATTTACATTTGAATTTATTATCATCAAGGTGTACAAATTCGCGGATCAA 1692
QY 247 TrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTyr 266
Db 1693 TGAATGTCAGATCAACACCATTCACCAATGTGAAGTCATCAGGTACT-----TAT 1743
QY 267 TyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyrArg 286
```

```
Db 1744 TATGACGAGCTATTGGTATATATTGAAAAGTATATTAATTATGACCAGCAATACCAT 1803
QY 287 ThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTyr 306
Db 1804 AAGGATTGAATCACCTTTAAAGAAATCAGAAAAAATCACATGGATGCTTTATAACACATAT 1863
QY 307 ArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIle 326
Db 1864 CGTCGAGAAATGACCTTTAAATTGTATTGGATCTGTGCGCAACTTTTCCCTTTTATGATATA 1923
QY 327 LysArgTyrArgAspSerIleGlyGlyLeuGluValLysGlyIleLysAsnGluLeuThr 346
Db 1924 CGTCGTITTT-----CCAGAGGAGTGAAGAACTAGAAITTAACA 1959
QY 347 ArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeuArgValGlnPro 366
Db 1960 AGAGAGGTTTATACAAGTTTAGATCATTTAAACACGACCACCA----- 2001
QY 367 AsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeu 386
Db 2002 -----GGGCTATTACTTTGGCTG 2019
QY 387 GluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIleSer 406
Db 2020 TCAGATATTGAGTTATACACGAGAGTGTGCAGAAAGCGGATTTATTATCAGGTATT--- 2076
QY 407 AsnArgAspAlaProThrTyrSerAsnThr-----IleThrGluThrLeuTyrGly 423
Db 2077 ---CGAGAGTCTAAATATATATCTGGTAAATCAATTTTTTACGATGAAAAATATTATGTT 2133
QY 424 GluArgThrGly---SerProThrThrLysThrIleArgProPheGlySerTyrLysVal 442
Db 2134 AATACAAATAGATTAAAGTAAGCAGCTCATTTACATTTATTACCAGGCGAA---TTATGACT 2190
QY 443 SerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIleIle 462
Db 2191 CACTTAAGCATAAACCGCTCTTTTCAACAATAGCTGTGTATATAAATAAGTTATACAGTTTA 2250
QY 463 AsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsnThrLeuLysTyrSerAlaGly 482
Db 2251 ATTCAAAAATCGTATTTCACAACTTTTAAAAACGATAAT-----GAATATCAAAAAAT 2304
QY 483 GlySerLeuSerAsnTyr-----GlnAsnThrThrPheGlnPheProArgLys 499
Db 2305 TTTAATGTGAATATCAAAATGAACCTCAAGAAACTACA----- 2343
QY 500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHis 519
Db 2344 -----AACTATCTCTAATGATTATGTTGTTTCA---AACAGCCAAAAATTCAAACAT 2391
QY 520 IleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnIle 539
Db 2392 AATTATCTCATTTTCCATTAAATCATCCAC-----AAGTTAGAGTTT 2433
QY 540 LeuAsp-----ThrGlyValLeuGlyTrpThrHisSerSerValAspArgTyr 555
Db 2434 GCTGAGTATTTTCACTCTATATTTCATTTGCGACACACAATAGTGTAACTCCCAA 2493
QY 556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp 575
Db 2494 AATTAAATATCAGAAAGTGTGAGTACACAAATCCCATTTGGTAAAGCTTTACGAAGTT--- 2550
QY 576 ThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln 595
Db 2551 ACTAAACATTCAGTTATTAAGAGGACCGAGTTTACAGGTGGAGATTTAATGAACCTTCT 2610
QY 596 SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIle 615
Db 2611 GAT-----AATGTCTTATTAAATGTAAGCT---AGTCTTTTAAAAAATACGCTATA 2661
QY 616 ArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIle 635
```



```
Db 2662 AGTCTATTTTATGCTGCAAAATAACGCAATAGCTGTATCAATAGACGCTAGGTGATTCGGA 2721
Qy 636 ProGlyValIleGlyleProProGlnArgLeuAsnThrPheSerGlyThrAsnTyr 655
Db 2722 GCAGGAGTTCTA-----TTGCAACCTACCTTTCTAGAAAAGGAAC 2763
Qy 656 AsnAsn-----LeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSer 670
Db 2764 AATAATTTTACAATTCAGACCTTAATACTATAAGGATTTCAATATCATACACATTTTAGTT 2823
Qy 671 ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAsp----- 688
Db 2824 GATATTGAATTAACCCAGAGTGAGAAATTCATATCCATTTGAAGCAGAGGATGATTAT 2883
Qy 689 ValSerAsnSerIleLeuIleAspLysIleGluPheIleProIleThrSerSerMet 708
Db 2884 GAGGAGGAGTGATCTTTTAATTTGATAAATTAGAGTTCAAACTATAGATGAATAATTAT 2943
Qy 709 HisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePhe 728
Db 2944 -----ACTAATGAATGAATTTAGAGAGGCAAGCAAGAGCAGTGAATGTATTATTT 2994
Qy 729 ThrAsnHisThrLysThrLeu 735
Db 2995 ATAAACGCAACAAACGCTTTG 3015

RESULT 6
US-10-342-821-3
; Sequence 3, Application US/10342821
; Publication No. US20030154510A1
; GENERAL INFORMATION:
; APPLICANT: Jana, Wojciechowska
; APPLICANT: Evgeny, Lewitin
; APPLICANT: Ludmila, Revina
; APPLICANT: Igor, Zalumin
; APPLICANT: Galina, Chestukhina
; TITLE OF INVENTION: NOVEL DELTA-ENDOTOXINS AND NUCLEIC ACID SEQUENCES CODING THEREFOR
; FILE REFERENCE: S-30913C
; CURRENT APPLICATION NUMBER: US/10/342,821
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US 60/175,158
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 3
; LENGTH: 4896
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(4896)
; OTHER INFORMATION: Bacillus thuringiensis supsp. finitimus strain VKPM B-1161 (i axon)
; OTHER INFORMATION: :29337
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1129)..(4458)
; OTHER INFORMATION: product: Cry28Aa1 delta-endotoxin
; PUBLICATION INFORMATION:
; AUTHORS: Wojciechowska, et al.
; TITLE: Two novel delta-endotoxin gene families cry26 and cry28 from Bacillus thuringiensis sep. finitimus
; JOURNAL: FEBS Lett.
; VOLUME: 453
; ISSUE: 12
; PAGES: 46-48
; DATE: 1999-06-18
US-10-342-821-3

Alignment Scores:
Pred. No.: 1.54e-79 Length: 4896
Score: 849.00 Matches: 229
Percent Similarity: 48.76% Conservative: 106
Best Local Similarity: 33.33% Mismatches: 248
```

```
Query Match: 21.94% Indels: 104
DB: 6 Gaps: 25
US-10-783-417-2 (1-735) x US-10-342-821-3 (1-4896)
Qy 68 SerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAla 87
Db 1210 AGCAGTGATACATCGCTGTAGTAAGCGAGGAGTGTGGTGGTACTATATCTGACA 1269
Qy 88 GlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIle 107
Db 1270 -----GCCTTTCATCATTTGTTAATCCA-----GGTGTGCTACTTATA 1308
Qy 108 SerPheGlyThrIleThrValPheThrProAlaGlyGluGlnAsp---LysThrVal 126
Db 1309 TCATTTGGAACCTTGGCTCCGCTCTTTGGCTGATCCAGAGGAAGATCCAAAAAATTT 1368
Qy 127 ThrThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeuThrGluSerIle 146
Db 1369 TGTGTCAATTTATGAAACACGAGAGACCTTTTAAATCAACAATTTCTACAGCTGTA 1428
Qy 147 LysGlnLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGlnSerTyrAsnThr 166
Db 1429 AAGAATAAGCATAGCTCATCTAATGCTTTTAAAGATGATTATTAAGTACTATGAAGA 1488
Qy 167 AlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeuProSerSer 186
Db 1489 GCATTTAATGATTGGAAGAGA-----AATCCAAGTGCA 1521
Qy 187 AlaLeuGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPhe 206
Db 1522 -----AATACTGCCAGATTGGTATCACAGAGATTGAAAAACGCTCATTTCAATTTT 1572
Qy 207 IleArgGluIleProGlyPheGlnLeuGluThrLysThrLeuLeuLeuProIleTyr 226
Db 1573 GTAAGCAATGATGCCAACCTCCACCTTCCACGATGACACATTTATTATTAAGTTGCTAT 1632
Qy 227 AlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGlyAlaGluLeuAlaAspGlu 246
Db 1633 ACAGAAAGCTGCAAAATTTACATTTTGAATTTTACATCAAGGTGTACAATTCGCGGATCAA 1692
Qy 247 TrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTyr 266
Db 1693 TGGAAATGCAGATCAACCAATTCACCAATGTTGAAGTCATCAGGTACT-----TAT 1743
Qy 267 TyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyrArg 286
Db 1744 TATGACGAGCTATTGGTATATATTGAAAGATATATTATTATTCACCAAGACATACCAT 1803
Qy 287 ThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTyr 306
Db 1804 AAAGGATTGAATCACCTTAAGAAATCAGAAAAATCACATGGGATGCTTATAACACATAT 1863
Qy 307 ArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIle 326
Db 1864 CGTCGAGAAATGACCTTAATTTGATTTGATCTGTGCAACTTTTCCTTTTATCATATA 1923
Qy 327 LysArgTyrArgAspSerIleGlyGlyLeuValLysGlyIleLysAsnGluLeuThr 346
Db 1924 CGTCGTTTT-----CCAAAGAGAGGTAGAACTAGAAATTAACA 1959
Qy 347 ArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeuArgValGlnPro 366
Db 1960 AGAGAGTTTATACAGTTTATAGATCATTTAACAGCACCA----- 2001
Qy 367 AsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeu 386
Db 2002 -----GGGCTATTTTACTTGGCTG 2019
Qy 387 GluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIleSer 406
Db 2020 TCAGATATTGAGTTATACACGAGAGTGTGGCAGAGCGGATTTATTCAGGTATT--- 2076
```

```
QY 407 AsnArgAspAlaProThrTyrSerAsnThr-----lleThrGluThrLeuTyrGly 423
Db 2077 ---CGAGAGCTAAATATATATATCGGTAACTATTTTACGATGAAATAATATTATGCT 2133
QY 424 GluArgThrGly---SerProThrThrLysThrLysThrLysThrLysVal 442
Db 2134 AATACAAATAGATTAAGTAAGCAGCTCATTTACATTTATACCGCGAA---TTTATGACT 2190
QY 443 SerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIle 462
Db 2191 CACTTAAGCATAAACCGCTCTTTTCAACCAATAGCTGGTATATAATAGTTATACAGTTTA 2250
QY 463 AsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsnThrLeuLysTyrSerAlaGly 482
Db 2251 ATTCAAAAATCGTATTCACAACTTTTAAACAGATAAT-----GAATATCAAAAAAT 2304
QY 483 GlySerLeuSerAsnTyr-----GlnAsnThrThrPhePheGlnPheProArgLys 499
Db 2305 TTTTAATGTGAATCAAAATGAACCTCAAGAACTACA----- 2343
QY 500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHis 519
Db 2344 -----AATATCTCTAATATGATTTATGTTGTTCA---AACAGCCAAAATTCACAACT 2391
QY 520 IleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuIle 539
Db 2392 AATTATATCTCATTTTCCATTAATCATCCAC-----AAGTTAGAGTTT 2433
QY 540 LeuAsp-----ThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyr 555
Db 2434 GCTGAGTATTTTCACTCTATATTTGCTATTTAGTTGGACACACATATGTTAACTCCAA 2493
QY 556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp 575
Db 2494 AATTAAATATCAGAAAGTGAGTACACAAATCCATTCCTGTTAAAGCTTACGAAAGTT-- 2550
QY 576 ThrAsnSerLysValIleGluGlyProGlyHisThrGlyAsnLeuValTyrLeuGln 595
Db 2551 ACTAAACATTCAGTTATTAAGAGGACCAAGGTTTATACAGTGGAGATTAAATAGAACTTCGT 2610
QY 596 SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIle 615
Db 2611 GAT-----AAATGTTCTATTAATGTTAAAGCT---AGTCTTTTAAATAATACGCTATA 2661
QY 616 ArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIle 635
Db 2662 AGTCTATTTATGCTGCATAATACGCAATAGCTGTATCAATAGACGTAGGTATCCGGA 2721
QY 636 ProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyr 655
Db 2722 GCAGGAGTTCTA-----TTGCAACCTACCTTTTCTAGAAAAGGGAAC 2763
QY 656 AsnAsn-----LeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSer 670
Db 2764 AATAATTTTACAATTCAGACCTTAATATAAGGATTTTCAATATCATACACTTTTATGTT 2823
QY 671 ThrValThrProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAsp----- 688
Db 2824 GATATTGAATATCCGNAAGTGAAGAAATTCATATCCATTTGAAGCAGGAGATGATAT 2883
QY 689 ValSerAsnSerIleLeuIleAspLysIleGluPheIleProIleThrSerSerMet 708
Db 2884 GAGGAGGAGTATCTTTTAAATGATAATTTAGATTTCAACCTATAGATGAATAATAT 2943
QY 709 HisGlnAsnArgGlnLysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePhe 728
Db 2944 -----ACTAATGAATAATTTAGAGAAGGCAAGAGCAGTGAATGTTATTTT 2994
QY 729 ThrAsnHisThrLysThrLeu 735
Db 2995 ATAAACGCAACAAACGCTTTG 3015
```

RESULT 7

```
US-10-929-754-2
; Sequence 2, Application US/10929754
; Publication No. US20050124803A1
; GENERAL INFORMATION:
; APPLICANT: ABDULLAH, MODH AMIR
; TITLE OF INVENTION: INSECTICIDAL CRYSTAL PROTEINS WITH ENHANCED TOXICITY
; FILE REFERENCE: 22727/04179
; CURRENT APPLICATION NUMBER: US/10/929,754
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 60/498,826
; FILING DATE: 2003-08-29
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 2
; LENGTH: 3684
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-10-929-754-2
```

```
Alignment Scores:
Pred. No.: 4,93e-70 Length: 3684
Score: 759.00 Matches: 234
Percent Similarity: 47.40% Conservative: 122
Best Local Similarity: 31.16% Mismatches: 267
Query Match: 19.62% Indels: 128
DB: 9 Gaps: 34
```

US-10-783-417-2 (1-735) x US-10-929-754-2 (1-3684)

```
QY 23 AsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnProLeuGln 42
Db 154 AATATGAATTCAGCG-----TATCCGTTAGCGAATGACTTACAAGGGTCAATGAAA 204
QY 43 AsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGlyAspAsn 62
Db 205 AACACAGCACTATAAAGATTGGCTAGCCATGTGGAATAACCAACAGTATGCGGTTAT 264
QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSer 82
Db 265 -----CCAGCTGCGATTAATTCCTTCTCAGTTAGTACCGCTTTTAAAGTAGCT 312
QY 83 GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
Db 313 GGAGCTATC-----CTTAAATTTGTAAACCCACCTCGCAGGTACTGTC 354
QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGly---Glu 121
Db 355 TTAACCGTACTT-----AGCGCGTGTCTCTATTTCTTGGCCGCACTAATCTCCA 405
QY 122 GlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThrPro 141
Db 406 ACGCTGAAGAGTTTGGATGATTTTCATGACCAATACAGGAATCTTATGATCAAACT 465
QY 142 LeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeu 161
Db 466 GTAACAGCTTATGTACGAACAGATGCAAAATGCAAAATGACGTTGTGAAAGATATTATA 525
QY 162 GlnSerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaProGly 181
Db 526 GATCAATATACAACTAAATTTTACACTTGGAAAGA----- 561
QY 182 LeuProProSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGluAsn 201
Db 562 GAGCTAATAACCACTGCTATAGAACAGCAGTAATACT-----CAATTT---AAC 609
QY 202 ValHisAspPheIleArgGluIleProGlyPhe-----GlnLeuGluThrTyrLys 219
Db 610 TTAACCACTGCGCAAACTTCGAGAGACCGCAGTTTATTTTAGCAACTTAGTGGTTATGAA 669
QY 220 ThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGln 239
Db 670 TTATTGTTATTACCAATATACGCAAGTAGCAAAATTTCAATTTTCTTTTAAAGAGAT 729
```

QY 240 GlyAlaGluLeuAlaAspGluThrProAsnAlaAspIleHisProSerGluIleGluProAsn 259
DB 730 GGCCTCATAAATGCAAGAATGGTCT- - - - -TTAGCAGCTAGT 768
QY 260 AlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsn 279
DB 769 GCTGGT- - - - -GACCACTATATAACACTATGTCGTCAGTACACTAAAGAAATATATTGCA 822
QY 280 TyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLys 299
DB 823 CATAGCATTCACGTGTATATAAAGTTAGATGCTACTTAGAATAATAATCTAATGACAA 882
QY 300 TrpSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSer 319
DB 883 TGGATTACGTTTAAATGATTATAAAGAGAGATGACTATTCAAGTATTAGATATACGCT 942
QY 320 GlnPheSerLeuTyrAspIleLysArgTyr- - -ArgAspSerIleGlyGlyIleGluVal 338
DB 943 CTTTTCGCGAGTTATGATCCAGCTGATACCCCTGCGGACAAATAGATATACGAACTA 1002
QY 339 LysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArg 358
DB 1003 TCA- - - - -AAAACAGATTACAGAGAGATTATACAGCT- - - - - 1038
QY 359 LeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAla 378
DB 1039 TTAGTAGAATCTCTCTAGTAAATCTATAGCAGCACTGGAGGCGACCTTACACGA- - - 1095
QY 379 SerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsnPhe 398
DB 1096 GATGTTTCATTATTCATCTGGCTAAAGAGAGATGATTTCTGGACC- - -AATACTATATAT 1152
QY 399 GlyAsnArgLeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThr 418
DB 1153 CAAGATTAAAGATTTTATCTGCTCAATAAAATGGGTTTTCATATACAAATCTCTGCA 1212
QY 419 - - - - -GluThrLeuTyrGlyGluArg- - -ThrGlySerProThrThrLysThrIle 434
DB 1213 ATGCAAGAAAGTGGATTTATGGAAGTCTGGTTTGGTTTCAATCTTACTCATCAAT 1272
QY 435 ArgProPheGluSer- - -TyrLysValSerIleValThrAspArgGlnSerProVal 453
DB 1273 CAACCTTAATCTAATGTTTATAAACTTCTATC- - -ACAGATAGTCTGCC- - - - - 1323
QY 454 SerProIleGlnProHisPheIleLeuAsnGluLeuTyrLeuAsnGlySerSer 473
DB 1324 - - - - - 1326
QY 474 AsnAsnThrLeuLys- - - - -TyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsn 490
DB 1327 AATCGAGTTACAAAAATGATTTCTACAAAATTTGATGGTACTCTTGCTCTTATATCA 1386
QY 491 - - - - -ThrThrPheGlnPheProArgLysLys 500
DB 1387 AATATAACCAACTCTCTGAAGTTTAAAGGACACACATTTTGGATTTTCAACAAATGAG 1446
QY 501 AspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIle 520
DB 1447 AAC- - - - -ACACCTAAT- - -CAACCAACTGTAATGATTATACGCATATT 1488
QY 521 LeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuIleLeu 540
DB 1489 TTAAGCTAT- - - - -ATAAAACTGATGTATATA 1515
QY 541 AspThrGly- - - - -ValLeuGlyTyrThrHisSerSerValAspArgTyrAsn 556
DB 1516 GATTATAACAGTAACAGGGTTTCATTGCTTGACACATAAGATTGTTGACCTTAATAAT 1575
QY 557 AlaIleSerAspLysIleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThr 576
DB 1576 CAAATATACACAGATGCTATCACACAAGTCCGCGCGTAAAAATCTAATCTTCTTGAATGCA 1635

QY 577 AsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSer 596
DB 1636 ACAGTAAAGTAATCAAGGACCTGGTCATACAGGGGGGATCTAGTCTCTTCAACAGC 1695
QY 597 - - - - -GlnGlyArgLeuGluIleThrCysGluThrPro- - - - -AsnSer 609
DB 1696 AATGGTACTCTATCAGGCAAGATGGAGATTCAATGTAACCAAGTATTTTAAATGATCCT 1755
QY 610 ThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuPro 629
DB 1756 ACAAGAGTTACGGATTACGCATACAGTTATGCTGCTCAATAGTCCA- - - - -ATTGTATTG 1809
QY 630 AsnIleSerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsnAsnThr 649
DB 1810 AATGTATCATATGATTACAAAGGAGTTTCTAGACAGAACACGATTAGTACAGAAATCTACG 1869
QY 650 PheSerGlyThrAsnTyr- - - - -AsnAsnLeuGlnTyrGlyAspPheGlyTyrPhe 666
DB 1870 TTTTCAAGACTAATAATAATAATACCTACAGATTAAATAATGAGAGTTTAGATACAAA 1929
QY 667 Gln- - - - -PheProSerThrValThrLeuProLeuAsnArgAsn- - - - -IleProPheIle 683
DB 1930 GATCCITTTGATGCAATTTGACCGATGAGATTATCTTCTAATCACTGATACTATAGCT 1989
QY 684 PheAsnArgAlaAspVal- - -SerAsnSerIleLeuIleIleAspLysIleGluPheIle 702
DB 1990 ATTCAACCAATTAACATGACTTCAATAATAATCAAGTGATTATTGACAGAAATCGAAATTT 2049
QY 703 ProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThr 722
DB 2050 CCAATCACTCAATCTGTATTAGTAGAGACAGAACCAAAATTTAGAAATCAGAACGAGAA 2109
QY 723 LysIleAsnThrPhePheThrAsnHisThrLys 733
DB 2110 GTTGTGATGCAGCTGTTTACAAATGACGCGNAA 2142
RESULT 8
US-10-929-754-4
; Sequence 4, Application US/10929754
; Publication No. US20050124803A1
; GENERAL INFORMATION:
; APPLICANT: ABDULLAH, MODH AMIR
; APPLICANT: DEAN, DONALD H.
; TITLE OF INVENTION: INSECTICIDAL CRYSTAL PROTEINS WITH ENHANCED TOXICITY
; FILE REFERENCE: 22727/04179
; CURRENT APPLICATION NUMBER: US/10/929,754
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 60/498,826
; PRIOR FILING DATE: 2003-08-29
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 4
; LENGTH: 4391
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-10-929-754-4
Alignment Scores:
Pred. No.: 4,48e-64 Length: 4391
Score: 705.00 Matches: 225
Percent Similarity: 43.38% Conservative: 119
Best Local Similarity: 28.37% Mismatches: 278
Query Match: 18.22% Indels: 171
DB: 9 Gaps: 29
US-10-783-417-2 (1-735) x US-10-929-754-4 (1-4391)
QY 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
DB 719 ATGCATTTATTTGGGAATAGGAATGAATGACATATTA- - - - - 757
QY 21 PheProAsnArgAsnSerAsnAspSer- - - - -ArgTyrProTyr 33
DB 21 PheProAsnArgAsnSerAsnAspSer- - - - -ArgTyrProTyr 33

758	-----AATGCTTCATCAAAAGATTCAAAATGTCTAATACTTATCGAGGATATCGGTTA	811
Qy	34 ThrAsnAsnProAsnGlnProLeuGlnAsnThrAsnTyrIysGlnTyrLeuAsnMetCys	53
Db	812 GCAATCCACAACAAGATTTAATGCCAAAATACAAATTTATAAGATTGGCTTAATGATGT	871
Qy	54 GlnGlyAsnThrGlnTyrGlyAspAsnPheGluThrPheAlaSerAlaAspThrIleAla	73
Db	872 GAAGGGGTAT-----CATATAGAAAATCCTAGAGAAGCAAGCGTTAGAGCT	916
Qy	74 AlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThr	93
Db	917 GGACTTCGGTAAGGATTAGGTATAGTATAGTACAACTGAGGTTCTTTGGTGGT---TCT	973
Qy	94 SerIleSerGlyProIleGlyIleIleGlyAlaIleIleSerPheGlyThrLeuIle	113
Db	974 ATTATCTTAGATACAATGGATTG-----TTTTACCAAAATTCA	1012
Qy	114 ThrValPheTrpProAlaGlyGluGlnAspIysThrValTrpThrGlnPheIleLysMet	133
Db	1013 GAGCTACTTTGGCCAGAGGATGATACCAGCAATACATCTTGGCAAGATATATTGAATCAT	1072
Qy	134 GlyGluIlePheValAspThrProLeuThrGlnSerIleIysGlnLeuIysLeuGlnThr	153
Db	1073 GTAGAGATCTTATAGACAAACGAACTAATGAGGTTATACGAGAAATGCAATTAGAAACA	1132
Qy	154 LeuGluGlyPheArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrPArgLys	173
Db	1133 TTAGCAGATTACAGGGTAAGTTGATGATATTAACAATTTGGTTGAAGAANTGGAAGAC	1192
Qy	174 LeuLysArgLeuGlnAlaProGlyLeuProProSerSerAlaLeuGlnAlaLeuLeu	193
Db	1193 -----GATCCAAAATCTACAGTAATTTAAGC-----	1219
Qy	194 ThrLeuLysIleArgPheGluAsnValHisAsnAspPhe-----IleArgGluIle	210
Db	1220 ACCTTAGTAACCAAGTTTACGGCTCTTGATTTCAGATTTAATGGTGCTATAAGGACAGTT	1279
Qy	211 -----ProGlyPheGlnLeuGluThrTyrIysThrLeuLeuLeuProIle	225
Db	1280 AATAATCAGGGAGTCCAGGTATAGGTTA-----CTTTATTGCCCTGTC	1324
Qy	226 TyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlyAlaGluLeuAlaAsp	245
Db	1325 TATGCACAAATTCGAATCTGCATTTACTTTTATACGGGATGCTCAGATTTATGGAGAT	1384
Qy	246 GluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAsp	265
Db	1385 AAATGG-----TGGAGCGCACAGCATATGCTCGTGATAAT	1420
Qy	266 TyrTyrIysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyr	285
Db	1421 TATTACCAATATACNATTAGAGAAACAAAGGAATATACAGAAATTTGTATAAATTTGGTAT	1480
Qy	286 ArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePheAsnAsp	305
Db	1481 AATAAGGGTTTAAATGATTTTAGAACACAGCAGT-----CAATGGGTAAACTTTATCGT	1534
Qy	306 TyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyrAsp	325
Db	1535 TATCGTAGAGAAATGACTCTTACTGTATTAGATATTTTCAATGTGTCCCTATTATTATGAC	1594
Qy	326 IleLysArgTyrArgAspSerIleGlyIleGluValIysGlyIleLysAsnGluLeu	345
Db	1595 GCGAGATTATCTCTACAGAA-----GTAATAAACCGAACA	1630
Qy	346 ThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeuArgValGln	365
Db	1631 ACTAGGGAATTTATTCAGATGTATTAAATGGGAGATATATGCACTTATGACTCTCTAT	1690
Qy	366 ProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPheLysLeuPheSerPhe	385
Db	1691 TTTTCTTTTTCAGAAAGCTGAATCACTTTATACAGGGCAACC---CATCTCTTCACATGG	1747

Qy	386	LeuGluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIle	405
Db	1748	CTAAAGGAAATTCGAATTTGTAACCAATTCATTCTCTATTGGACATTTTATCAGCTGGT	1807
Qy	406	SerAsnArgAspAlaProThrTyrSerAsnThrIleThrGlu-----	419
Db	1808	CAAAAYTAAGTATTCTTATACATAATAATCTAGATTAAACGAGGGCTCTTTTATAGGGACAC	1867
Qy	420	--ThrLeuTyrGlyGluArgThrGlySerProThrIleThrIleThrIleArgProPheGlu	438
Db	1868	GACACAGATTATGGT-----GGGACTCTCTCTACCAATTAATATT---CCATCAAAAT	1915
Qy	439	SerTyrIysValSerIleValThrAspArgGln-----SerProPro	452
Db	1916	TCGTATGTATATAATTTATGTCGCGGAAAAATTATGAATATATATTTATCTCTGGGTGATCCT	1975
Qy	453	ValSerProIleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySer	472
Db	1976	GTAATAATTACAAAAATGAATTTTCTGTAAACA-----GATAATAAT	2017
Qy	473	SerAsnAsnThrLeuIlystYrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThr	492
Db	2018	TCITCAAAAGAAATTAATTTATGTGTGCACACAGAACGAATAACCTGTTGTCGACAGAT	2077
Qy	493	PhePheGlnPheProArgIlystYsAspCysAsnLeuValIleAspProGlyCysSerPro	512
Db	2078	TTTGATTTTCTCACTAATAAGAGGGAACGTAGCTTAGCA-----	2116
Qy	513	AsnPheAsnAsnTyrSerHisIleLeuSerHisPheSerLeu-----	526
Db	2117	AAATATAATGATTATATATCATATTTTATCTCTATATGTTAATTAATGGGAAACGTTTGGT	2176
Qy	527	-----PheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuaspThrGly	543
Db	2177	CAGAAACGTCAATCGTTATCGTTT-----	2200
Qy	544	ValLeuGlyTprThrHisSerSerValAspArgTyrAsnAlaIleSerAspIlyIleIle	563
Db	2201	-----GCTTTTACACATAGTAGTGTGATCTCTAATAATACCATTCGACGAATAAAAT	2254
Qy	564	ThrMetIleProAlaIleIysGlyAsnAsnLeuAspThrAsnSerIysValIleGluGly	583
Db	2255	ACGCAAAATTCCTGTAGTGAAGAGCTCGAGTATAAATGGATCGATTTCAATTGGAAGAAGGT	2314
Qy	584	ProGlyHisIthrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIleThr	603
Db	2315	CCCGGATTTACGGGAGGAGATTTGGTAAAGATGAGAGCAGATTCAGGTTTAACTATCGCT	2374
Qy	604	CysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn---	622
Db	2375	TTTAAAGCTGAATTATAGATAAAAAATATCGTGTTCGAATAGATTATAAATGTAAC	2434
Qy	623	-----GlyValaGly-----AsnThrLeuPro	629
Db	2435	AGTTCTAAATTAATACTACGAAAAATGMAAGGGGAGGTTATATACAACAACAAAATTCAC	2494
Qy	630	AsnIleSerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsnAsnThr	649
Db	2495	AAATTTCTCCCA-----	2509
Qy	650	PheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGln---Phe	668
Db	2510	-----TATGGAGCCCTTTCTTATTATTAGAGTCTTTT	2539
Qy	669	ProSerThrValThr-----LeuProLeuAsnArgAsnIleProPheIlePheAsn	685
Db	2540	ACTATAACTACGACAGAAAAATATATTGATTTCACAATGGAGGTAACATATCCGTATGGT	2599
Qy	686	ArgAlaAspValSerAsn---SerIleLeuIleIleAspIlyIleGluPheIlePro-I	704
Db	2600	AGACAGTTTCTTGAAGATATACCATCTTATATAGATAAAATCGAATTCCTCCCAACT	2659

Qy	462	IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys	478
Db	1291	ATGCCAAGAGTCGAGT-----CATCGTTAAACCAATTTGAATATATACCGAAGACGTTAAAG	1350
Qy	479	TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr	491
Db	1351	TATTAATCGGTTTCCAAGATATTATAGCGGG-----	1383
Qy	492	ThrPhePheGlnPheProArgLeuAspCysAsnLeuValIleAepProGlyCysSer	511
Db	1384	-----ACAAGAGATTCGGAAATTAGAAATTACCTCCAGAAACTTCA	1422
Qy	512	-----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPhe-----SerLeuPheThr	528
Db	1423	GATCAACCAAAATTTATGATCATATAGCCATAGATTATGTCTATATACAAAGTATTTCCGCG	1482
Qy	529	TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThr	548
Db	1483	ACGGTTCAACTACCGGATTA-----GTACTGTATTTTCTTGACA	1524
Qy	549	HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla	568
Db	1525	CATCGGAGTCCGATCTTATAAATGCAGTTCATTACAGATAAAATTAATCAGATTCGGTC	1584
Qy	569	IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu	582
Db	1585	GTAAAGGTTTCTGATTTGGCTCCCTCTATAACAGAGGGCCAAATAATACCGTTGTATCG	1644
Qy	583	GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGly-----	598
Db	1645	GGTCTCGATTTACAGGGGGGGATATAAAGTAATAAGAAATGAGTAATATATATCA	1704
Qy	599	-----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg	616
Db	1705	CATATCGTGTTAAANITTA-----GACATTAACAAGAAATATAGTATGAGG	1752
Qy	617	LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro	636
Db	1753	ATTCGGTATGCTTC-----GCTAATAATACTGAATTTTATATAAATCCTTCTCGAAGAA	1806
Qy	637	GlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn	656
Db	1807	AACGTTAAATCT---CACGGTCTAAAAAATCATGAATAGAGGTGAAGCTTTTAAACATATAAT	1863
Qy	657	AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValThr	673
Db	1864	AAATTTAAATATGCGACTTTGCCCCCTATTAAATTTACGAACCGAACCTTTTCATTACT	1923
Qy	674	LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle	693
Db	1924	CTAGGGGCT-----ATATTTGAACGGGAAGACTTTCTTGGAAATTGAA	1965
Qy	694	LeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGlu	713
Db	1966	GCTTATATAGACCAATCGAATTTATCCAGTAGATGAGACATAT-----GAAGCG	2016
Qy	714	LysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsn	730
Db	2017	GAAACAAGATTAGAACGAGCGAAGAAAGACGATGTAATGCTTTGTTTACGAAT	2067

```

RESULT 10
US-10-414-637-3
; Sequence 3, Application US/10414637
; Publication No. US20030177528A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Nicholas B. Duck
; APPLICANT: Xiang Feng
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Theodore W. Kahn
; APPLICANT: Lynn E. Sims
; TITLE OF INVENTION: Genes Encoding No. US20030177528A1e1 Proteins With
; TITLE OF INVENTION: Pesticidal Activity Against Coleopterans

```

```

; FILE REFERENCE: 35718/237005
; CURRENT APPLICATION NUMBER: US/10/414,637
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: US/10/032,717
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: fastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3633
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3633)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Cyl1218-2
US-10-414-637-3

```

Alignment Scores:		
Pred. No.:	5.34e-64	3633
Score:	703.00	Matches: 219
Percent Similarity:	44.53%	Conservative: 127
Best Local Similarity:	28.19%	Mismatches: 289
Query Match:	18.17%	Indels: 142
DB:	6	Gaps: 32

US-10-783-417-2 (1-735) x US-10-414-637-3 (1-3633)

Qy	5	AsnAspAsnAsnGlnTyrGluIleAlaSerHisThrSerProTyrPheProAsnArg	24
Db	10	ATAAATCAAAATGAATGAATGAATATATAGATCGCACACTTCT	57
Qy	25	AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn	43
Db	58	GTATCCAAATGATTTACAGATACCCCTTTCCGAAATGAGCCAAACAAATGCGCTACAAAAT	117
Qy	44	ThrAsnTyrLysGluTrrpLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn	62
Db	118	ATGGATTATAAAGATTATTTAAAAATGTCCTCGGGAATGCTAGTAGAATACCCTGGTTCA	177
Qy	63	PheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIleValSer	82
Db	178	CCTGAAGTACTTGTTCGCGACACAGAT-----GCAGCTAAGGCGCAATTGATATAGTA	231
Qy	83	GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIle	102
Db	232	GGTAATATTACTATCAGGTTTAGGG-----GTCCATTGTGT	267
Qy	103	GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrrpAlaGlyGluGln	122
Db	268	GGGCGGATAGTACTTCTTACTCAACTTATTGATATCTGTGGCCTTCAGGGCAA---	324
Qy	123	AspLysThrValTrrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu	142
Db	325	---AAGAGTCAATGGGAGATTTTATGGACCAAGTAGAAGAACTCATAAATCAAAAATA	381
Qy	143	ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln	162
Db	382	GCAGATATGCAGGAATAAAGCGCTTTCGGAATTAGAAGGATTAGGTAAATAATTACCAA	441
Qy	163	SerTyrAsnThrAlaLeuAspAspTrrpArgLysLeuLysArgLeuGlnAlaProGlyLeu	182
Db	442	TTATATCTAATCGCGCTTGAAGATGAAAGAA-----	474
Qy	183	ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleAlaPheGluAsnVal	202
Db	475	AATCCAATGGTTCA-----AGAGCCCTTACAGATGCGCAAAATCGATTTGAAATCCTTG	528
Qy	203	HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluGluThrTyrLysThrLeuLeu	222


```
; LOCATION: (1)...(3633)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Cry1218-2
; US-10-606-320-3

Alignment Scores:
Pred. No.: 5,34e-64      Length: 3633
Score: 703.00           Matches: 219
Percent Similarity: 44.53%      Conservative: 127
Best Local Similarity: 28.19%    Mismatches: 289
Query Match: 18.17%            Indels: 142
DB: 7                        Gaps: 32

US-10-783-417-2 (1-735) x US-10-606-320-3 (1-3633)

QY 5 AsnAspAsnGluThrGluLeuValSerHisThrSerProTyrPheProAsnArg 24
DB 10 AATAATCAAAATGAATATGAATATATAGATCGACACTTCT-----ACTTCT 57

QY 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
DB 58 GTATCCAAATGATTTCTACAGATACCTTTTTCGGAATGAGCCCAAAATGCGCTACAAAAT 117

QY 44 ThrAsnTyrLysGluThrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
DB 118 ATGATTATTAAGATATTATTAATAAATGTCTGCGGGAATGCTAGTAATACCTCGGTTC 177

QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIleValSer 82
DB 178 CCTGAAGTACTTGTAGCGGACAAGAT-----GCAGCTAAGCGCGCAATGATATAGTA 231

QY 83 GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIle 102
DB 232 GGTAAATTACTATCAGGTTTAGGG-----GTCCCAATTGTT 267

QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGlyGluGln 122
DB 268 GGCCCGGATAGTAGTCTTTTATCTACTCAACTTATTGATATTCTGTGGCCTTCAGGGCAA--- 324

QY 123 AspLysThrValThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu 142
DB 325 ---AAGAGTCAATGGGAGATTTTATGGACAACAGTAGAAGACTCAATAAATCAAAAATA 381

QY 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
DB 382 GCAGATATCAAGGAATAAGCGCTTTCGGAATTAGAGGATTTAGTAATAATTACCAA 441

QY 163 SerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
DB 442 TTATATCTAACTGCGCTTGAAGAATGAAAGAA----- 474

QY 183 ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal 202
DB 475 AATCCAAATGGTTCA-----AGAGCCTACGAGATGTGCGAAATCGAATCGAATCCTG 528

QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
DB 529 GATAGTTATTATAGCAATACATGCCATCTTTTCGAGTGACAATTTTGAAGTACCAATTC 588

QY 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGlyAlaGlu 242
DB 589 CTTACAGTATATACAGCGCAGCAACCTTCTACTGTTATTAAAGACGCTTCAATT 648

QY 243 LeuAlaAspGluThrPheAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
DB 649 TTTGGAGAGAAATGGGATGG-----TCTCAACACCACT 681

QY 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
DB 682 ATTAATAACTATTATGATCTCAATGAAACTTACTGCAGAAATATTCTGATCACTGTGTA 741
```

```
QY 283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTyrSerIle 302
DB 742 AAGTGGTATGAACCTGGTTTAGCAAAATTAAGAGGACGAGCGCTTAAACAATGGGTCGAC 801

QY 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
DB 802 TATAACCAATTCGCTAGAGAAATGACACTGACGGTTTGTAGATGTTGCTGATATTATCCCA 861

QY 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIleLys 342
DB 862 AATTATGACACAGCCGCTAC-----CCAATGAAACGAAAGCA----- 900

QY 343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362
DB 901 ---CAACTAACAGGAAAGTATATACA-----GATCCACTGGGCGCGGTA 942

QY 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
DB 943 AACGTGTCTTCAATGGTTCCTGGTATGACAAAGCACCTTCTTTCGGAGTGATAGATCA 1002

QY 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
DB 1003 TCCGTATTTCGACCACCC---CATGTAATTGATTATATAACGGGACTCACAGTGATACA 1059

QY 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408
DB 1060 CAATCAAGAACCAATTTCTTCGCTGCTATATAAGACATTTGGGCTGGTCAATAAATG 1119

QY 409 AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySer 428
DB 1120 TATCATCGGATTTTAGTGAATAATATATAACAGATGTATGGAACATAATCAAAATCTA 1179

QY 429 ProThrThrIleArgProPheGluSerTyr-----Lys 441
DB 1180 CACAGCACTAGTACTCTT---GATTTTACGAATTATGATATTTTACAGACGTTATCAAAA 1236

QY 442 ValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIle 461
DB 1237 GATCGGTGCTCTGTGATATTGTTTCTCTGGTTATACGTATATA-----TTTTTGA 1290

QY 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
DB 1291 ATGCCAGAGTCGAGTGTTCATGTGTAACCAATTTGAATAATACCAGAAAGACGTTAAAG 1350

QY 479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr 491
DB 1351 TATAATCCGGTTTCCAAAGATATTATACGGGG----- 1383

QY 492 ThrPhePheGlnPheProArgLysAspCysAsnLeuValIleAspProGlyCysSer 511
DB 1384 -----ACAAGAGATTCGGAATTAGAAATTACCTCCAGAAACTTCA 1422

QY 512 -----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPhe---SerLeuPheThr 528
DB 1423 GATCAACCAAAATATAGTCTATAGCCATAGATTATGTCATATCAACAAGTATCCCGCG 1482

QY 529 TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyThr 548
DB 1483 ACGGGTCAACTACCGGAATTA-----GTACCTGTATTTTCTTGGACA 1524

QY 549 HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla 568
DB 1525 CATCGGAGTCCGATCTTATAAATGCGAGTTCATTTCAGATATAAAATTTACTCAGATTCCGGTC 1584

QY 569 IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu 582
DB 1585 GTAAAGGTTTCTGATTTGGCTCCCTCTATAACAGAGGCGCAAAATAATACCTGTATCG 1644

QY 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGly----- 598
DB 1645 GGTCTCGATTTTACAGGGGGGGGATATAAAGTAATAAGAAATAGAGTAATATATATCA 1704

QY 599 -----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616
```



```

Db 1705 CATATGGTGTAAATTTCA-...::: ||| ::|||
Qy 617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
Db 1753 ATTCCGTAGCTCC-...GCTAATAACTGAATTTTATATAAATCTCTCGAAGAA 1806
Qy 637 GlyValIleGlyLeuProGlnArgLeuAsnThrPheSerGlyThrAsnTyrAsn 656
Db 1807 AACGTTAAATCT--CACGCTCAAAAAAATGAAATGAGTGAAGCTTTTAAACATATAAT 1863
Qy 657 AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValThr 673
Db 1864 AAATTTAATATCGACTTTGCCCTTATTAATAATTACGACCAACCGAACCCTTTCATTACT 1923
Qy 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693
Db 1924 CTAGGGGCT-----ATATTTGAACCGGAAGACTTTCTTGGNAATTGAA 1965
Qy 694 LeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGlu 713
Db 1966 GCTTATATAGACCGAATCGAATTTATCCAGTAGATGAGACATAT-----GAAGCG 2016
Qy 714 LysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsn 730
Db 2017 GAACAAGATTAGAACGACGCAAGAAAGCAGTGAATGCCTTGTGTTACGAAT 2067

```

RESULT 12

```

US-10-746-914-3
; Sequence 3, Application US/10746914
; Publication No. US20040210963A1
; GENERAL INFORMATION:
; APPLICANT: Albert L. Lu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE REFERENCE: Activity
; CURRENT APPLICATION NUMBER: US/10/746,914
; CURRENT FILING DATE: 2003-12-24
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3633
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3633)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Cry1218-2
US-10-746-914-3

```

Alignment Scores:

```

Pred. No.: 5,34e-64 Length: 3633
Score: 703.00 Matches: 219
Percent Similarity: 44.53% Conservative: 127
Best Local Similarity: 28.19% Mismatches: 289
Query Match: 18.17% Indels: 142
DB: 8 Gaps: 32

```

US-10-783-417-2 (1-735) x US-10-746-914-3 (1-3633)

```

Qy 5 AsnAspAsnGluTyrGluIleIleAspSerHisThrSerProTyrPheProAsnArg 24
Db 10 AATAATCAAAATGAATGAATTAATAATGATGCGACACTTCT-----ACTTCT 57

```

```

Qy 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
Db 58 GTATCCAAATGATTAACAGATACCTTTTCGGAATGAGCCACAAATCGCGTACAAAAT 117
Qy 44 ThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
Db 118 ATGATTTATAAAGATTAATTTAAAAATGTCGCGGGAATGCTAGTAATACCTCGGTCA 177
Qy 63 PheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSer 82
Db 178 CCTGAAGTACTTGTAGCGGACAAGAT-----CGAGCTAAGCGCGCAATTGATATAGTA 231
Qy 83 GlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
Db 232 GGTAAATTAATCTATCAGGTTTAGG-----GTCCCAATTTGTT 267
Qy 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGlyGluGln 122
Db 268 GGGCCGATAGTAGTCTTTTATCTCAACTTATGATTTCTGTGGCCTTCAGGGCAA--- 324
Qy 123 AspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu 142
Db 325 ---AAGAGTCATGCGGAGATTTTATGGAACAAGTAGAAGAACTCATAAATCAAAAATA 381
Qy 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
Db 382 CGAATATCAAGGAATAAAGCGCTTTCGGAATTAGAGGATTAGGTAAATAATTACCAA 441
Qy 163 SerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
Db 442 TTATATCTAACTGCGCTTGAAGATGGAAGAA----- 474
Qy 183 ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal 202
Db 475 AATCCAAATGGTTCA-----AGAGCCTTACGAGATGCGGAATCGATTTGAAATCCTG 528
Qy 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluTyrThrLysThrLeuLeu 222
Db 529 GATAGTTTATTAGCGCAATACATGCCATCTTTTCGAGTGACAAATTTTGAAGTACCATTC 588
Qy 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGlu 242
Db 589 CTTACAGTATATACACAGCGCAGCAACCTTCATTACTGTTTATTAAAGAGCGCTTCAATT 648
Qy 243 LeuAlaAspGluTyrPheAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
Db 649 TTTGGAGAAGAAATGGGATGG-----TCTACAACCACT 681
Qy 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
Db 682 ATTAATAACTATTATGATCGTCAAAATGAAACTTACTGCAGAAATATTCTGATCACGTGTGA 741
Qy 283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle 302
Db 742 AAGTGGTATGAACCTGCTTTAGCAAAATTAAGAGCGACGAGCGGTAAACAATGGTGCAG 801
Qy 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db 802 TATAACCAATTCGCTAGAGAAATGACACTGACGTTTTAGATGTTGTTGCAATTATCCCA 861
Qy 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyIleGluValLysGlyIleLys 342
Db 862 AATTATGACACACCGCAGTAC-----CCAATGGAACCAAGAACCA----- 900
Qy 343 AsnGluLeuThrArgGluIleTyrThrGluIleAsnPheAspArgLeuProGlnLeu 362
Db 901 ---CAACTAACAAGGGAAGTATATACA-----GATCCACTGCGCGCGTA 942
Qy 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
Db 943 AACGTGTCTTCAATTTGGTTCTGTTATGACAAAGACCTTCTTTCTGGAGTAGTAAGAATCA 1002
Qy 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnPheIlePheTyrThr 393

```

```
Db 1003 TCGGTTATTCGACCAACC---CGTGATTTGATTATATAACGGGACTCACAGTGTATACA 1059
Qy 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408
Db 1060 CAATCAAGAAGCAATTCCTCGCTCGCTATATAAGACATTCGGCTGGTGCATCAATAAAGC 1119
Qy 409 AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySer 428
Db 1120 TATCATCGGATTTTGTAGTAAATATATAAACAAGATGTATGAACATAATCAAAATCTA 1179
Qy 429 ProThrThrLysThrIleArgProPheGluSerTyr-----Lys 441
Db 1180 CACAGCACTAGTACTTT---GATTTACGAATTATGATATTTACAGACGCTTATCAAAA 1236
Qy 442 ValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIle 461
Db 1237 GATCGGTGTCCTTGTATTTTCTCGCTGATATGTTTCTCGTTATACGTATATA-----TTTTTGA 1290
Qy 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
Db 1291 ATGCCAAGAGTCGAGTTTTCATCGTAAACCAATTTGAATAATACCAAGAAAGACGTTAAAG 1350
Qy 479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr 491
Db 1351 TATAATCCGGTTTCCAAAGATATATAGCGGG----- 1383
Qy 492 ThrPhePheGlnPheProArgLysAspCysAsnLeuValIleAspProGlyCysSer 511
Db 1384 -----ACAAGAGATTCGGAATTAGAAATTACCTCCAGAAACTTCA 1422
Qy 512 -----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPhe-----SerLeuPheThr 528
Db 1423 GATCAACCAAAATATGATGATATAGCATAGATATATGTCATATACAAAGTATTTCCCGCG 1482
Qy 529 TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThr 548
Db 1483 ACGGTTCAACTACCGGATTA-----GTACCTGATTTTCTTGGACA 1524
Qy 549 HisSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla 568
Db 1525 CATCGGAGTCCGATCTTATAAATGCGATTCATTCAGATAAAATTAATCAGATTCGGTCC 1584
Qy 569 IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu 582
Db 1585 GTAAAGGTTTCTGATTTGGCTCCCTCTATAACAGAGGGCCAAATAATACCGTTGTATCG 1644
Qy 583 GlyProGlyHisThrGlyAsnLeuValTyrLeuGlnSerGlnGly----- 598
Db 1645 GGTCTCGATTTACAGGGGGGGGATAAATAAGTAATAAGAAATGGAGTAATATATATCA 1704
Qy 599 -----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616
Db 1705 CATATGCGTGTAAATTTCA-----GACATTAACAAGAAATATAGTATGAGG 1752
Qy 617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
Db 1753 ATTGCGTATGCTTCC-----GCTAATAATACTGAATTTTATATAAATCCTTCTGAAGAA 1806
Qy 637 GlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn 656
Db 1807 AACGTTAAATCT---CACGCTCAAAAAAATATGAATAGAGTGAAGCTTTAAACATATAAT 1863
Qy 657 AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValThr 673
Db 1864 AAATTTAAATTCGACCTTTGCCCTTATTAATTTTACGACCAACCCCTTTCATTACT 1923
Qy 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693
Db 1924 CTAGGGCT-----ATATTGAAGCGGAGACTTCTTGGAAATTGAA 1965
Qy 694 LeuIleIleAspLysIleGluPheIleProIleThrSerMetHisGlnAsnArgGlu 713
Db 1965 ----- 713
```

```
Db 1966 GCTTATATAGACCGAATCGAATTTTATCCAGTAGATGAGACATAT-----GAAGCG 2016
Qy 714 LysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsn 730
Db 2017 GAACAAGATTTAGAACGACGACGAAAGAAAGAGTGAATGCTTGTACGAAT 2067

RESULT 13
US-10-032-717-28
; Sequence 28, Application US/10032717
; Publication No. US20020151709A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Nicholas B. Duck
; APPLICANT: Xiang Peng
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Theodore W. Kahn
; APPLICANT: Lynn E. Sims
; TITLE OF INVENTION: Genes Encoding No. US20020151709A1el Proteins With
; FILE REFERENCE: 35718/237005
; CURRENT APPLICATION NUMBER: US/10/032,717
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 6613
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Genomic Cry1218-2
US-10-032-717-28

Alignment Scores:
Pred. NO.: 1,47e-63 Length: 6613
Score: 703.00 Matches: 219
Percent Similarity: 44.53% Conservative: 127
Best Local Similarity: 28.19% Mismatches: 289
Query Match: 18.17% Indels: 142
DB: 32

US-10-783-417-2 (1-735) x US-10-032-717-28 (1-6613)
Qy 5 AsnAspAsnAsnGluTyrGluIleAspSerHisThrSerProTyrPheProAsnArg 24
Db 1263 AATAATCAAAATGAATATGAATAATATATAGATCGACACCTTCT-----ACTTCT 1310
Qy 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
Db 1311 GTATCCCAATGATCTTAACAGATACCTTTTCGGAATGAGCCAAACAAATGCGGTACAAAT 1370
Qy 44 ThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
Db 1371 ATGATTATAAAGATATTATTAATAATGCTCGGGAATGCTAGTGAATACCTCGTTCA 1430
Qy 63 PheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIleValSer 82
Db 1431 CCTGAAGTACTTGTGTAGCGGCAAGAT-----GCAGCTAAGGCGCAATTGATATAGTA 1484
Qy 83 GlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
Db 1485 GGTAAATTAATCTATCAGGTTTAGGG-----GTCCCATTTGTT 1520
Qy 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGlyGluGln 122
Db 1521 GGGCGGATAGTACTTATTACTCAACTTATGATATCTCTGCTTACGGGCA--- 1577
Qy 123 AspLysThrValTyrThrGlnPheIleLysMetGlyIlePheValAspThrProLeu 142
Db 1578 ---AAGAGTCAATGGGAGATTTTATGGAACAAGTAGAAGAACTCATAAATCAAAAAATA 1634
```

Qy	143	ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln	162
Db	1635	GCAGATATTCGAAGGAATAAAGCCCTTTTCGGAATTAGAGGATTAGTAATAATTACCAA	1694
Qy	163	SerTyrAsnThrAlaLeuAspAspTirArgLysLeuLysArgLeuGlnAlaProGlyLeu	182
Db	1695	TTATATCTAATCGCGCTTGAAGATGAAAGAA	1727
Qy	183	ProProSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal	202
Db	1728	AATCCAATGGTTCA-----AGAGCCTTACGAGATGTGCGAATCGATTTGAAATCCTG	1781
Qy	203	HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu	222
Db	1782	GATAGTTTATTTAGCCAATACATGCCATCTTTTCGATGACAAATTTTGAAGTACCATTC	1841
Qy	223	LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGlu	242
Db	1842	CTTACAGCTATATACACAGGACGCAACCTTCATTTACTGTGTATTAAGGACGCTCAAT	1901
Qy	243	LeuAlaAspGluTirPheAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr	262
Db	1902	TTTCGAGAGAGATGGGATGG-----TCTACACCACT	1934
Qy	263	SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla	282
Db	1935	ATTAAATAACTATTATGATCGTCAAAATGAACTACTGCAGAAATATTCTGATCACTGTGTA	1994
Qy	283	AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle	302
Db	1995	AAGTGGTATGAACCTGTGTTTACGAAAATATAAAGGCACGACGCGCTAAACATGGGTGCAC	2054
Qy	303	PheAsnAspTyrArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer	322
Db	2055	TATNACCAATTCGTAGAGAAATGACACTGACGCTTTTAGATGTTGTGCATTATTCCCA	2114
Qy	323	LeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyLeuValLysGlyIleLys	342
Db	2115	AATTATGACACACGACGCTAG-----CCAATGGAAACGAAGCA-----	2153
Qy	343	AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu	362
Db	2154	--CAACTAAACAGGGAAGTATATACA-----GATCCACTGGCGCGGTA	2195
Qy	363	ArgValGln-----ProAsnLeuAlaThrMetGluTyr	373
Db	2196	AACGTGTCTTCAATGGTTCCTGTGTATGACAAAGACCTTCTTTCGGAGTGATAGAAATCA	2255
Qy	374	AsnLeuThrArgAlaSerPheLysPheLysPheSerPheLeuGluGlnPheIlePheTyrThr	393
Db	2256	TCCGTATTTCGACACCC--CATGTATTGATTATATAACGGGACTCACAGTGTATACA	2312
Qy	394	GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg	408
Db	2313	CAATCAAGAAGCAATTCTTCGCTCGCTATATAAGACATTTGGGTGTCATCAATAAGC	2372
Qy	409	AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySer	428
Db	2373	TATCATCGGATTTTACTGTGATAATTATATAAACAGATGTATGGAACATAACAAATCTA	2432
Qy	429	ProThrThrLysThrIleArgProPheGluSerTyr-----Lys	441
Db	2433	CACAGCACTAGACTTTT-----GATTTTACGAAATATGATATTTTACAGACGTTATCAAA	2489
Qy	442	ValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIle	461
Db	2490	GATCGGTGCTCTTGATATATGTTTCTCTGTTTATACGTATATA-----TTTTTGGG	2543
Qy	462	IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerSerAsnAsn-----ThrLeuLys	478
Db	2544	ATGCCAAGTCGAGTTTTTTCATGTGTTAAACCAATTTGAATTAATACCGAAGACGTTAAAG	2603

Qy 479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr 491
Db 2604 TATAAATCCGGTTCCTCCAAAGATATATATAGCGGG----- 2636
Qy 492 ThrPheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSer 511
Db 2637 -----ACAGAGATTCGGAAATTAGAATTACTCTCGAANAATTCA 2675
Qy 512 -----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPhe---SerLeuPheThr 528
Db 2676 GATCAACCAAAATATGAGTCATATAGCCATAGATTATGTCATATACAGATATCCCGCG 2735
Qy 529 TyrSerTyrValIleGlyLeuGlnLeuLeuAspThrGlyValLeuGlyTyrThr 548
Db 2736 ACGGGTTCAACTACCGGATTA-----GTACCTGTATTTCTTCGACA 2777
Qy 549 HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla 568
Db 2778 CATCGGAGTCCCGATCTTATAAATGCAGTTTCATTCAGATAAAATTAATCAGATTCGGTC 2837
Qy 569 IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu 582
Db 2838 GTAAAGTTTCTGATTTGGTCCCTCTATACAGGAGGCCAAATTAATACCGTTGTATCG 2897
Qy 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGly----- 598
Db 2898 GGTCTTGATTTACAGGGGGGGGATAATAAAGTAATAAGAAATGGAGTAATTATATCA 2957
Qy 599 -----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616
Db 2958 CATATGCGTGTAAAAATTCA-----GACATTAACAAGAAATATAGTAGTAGG 3005
Qy 617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
Db 3006 ATTCGGTATGCTCC-----GCTAATACTAGTAATTTATATAAATCCTCTGAGAA 3059
Qy 637 GlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn 656
Db 3060 AACGTTAAATCT--CACGCTCAAAAACTATGAATAGAGTGAAGCTTTAAACATATAAT 3116
Qy 657 AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValThr 673
Db 3117 AAATTTAATTATCGACTTTGCCCTATTAAATTTACGACACCGAACCTTTCTATTACT 3176
Qy 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693
Db 3177 CTAGGGCT-----ATATTTGAAGCGGAGAGACTTCTTGGAAATTGAA 3218
Qy 694 LeuIleIleAspLysIleGluPheIleProIleThrSerSerMethHisGlnAsnArgGlu 713
Db 3219 GCTTATATAGCCGAATCGAATTTATCCAGTAGATGAGACATAT-----GAAGCG 3269
Qy 714 LysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsn 730
Db 3270 GAACAAGATTTAGAACGAGCGAAGAAAGCAGTGAATGCCTTGTGTTACGAAT 3320

RESULT 14
US-10-414-637-28
; Sequence 28, Application US/10414637
; Publication No. US20030177528A1
; GENERAL INFORMATION:
; APPLICANT: Nicholas B. Duck
; APPLICANT: Xiang Peng
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Theodore W. Kahn
; APPLICANT: Lynn E. Sims
; TITLE OF INVENTION: Genes Encoding No. US20030177528A1el Proteins With
; TITLE OF INVENTION: Pesticidal Activity Against Coleopterans
; FILE REFERENCE: 35718/237005
; CURRENT APPLICATION NUMBER: US/10/414, 637
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: US/10/032,717

```
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 6613
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Genomic Cyl1218-2
US-10-414-637-28

Alignment Scores:
Pred. No.: 1,47e-63 Length: 6613
Score: 703.00 Matches: 219
Percent Similarity: 44.53% Conservative: 127
Best Local Similarity: 28.19% Mismatches: 289
Query Match: 18.17% Indels: 142
DB: 32

US-10-783-417-2 (1-735) x US-10-414-637-28 (1-6613)

QY 5 AsnAspAsnGluTyrGluIleAspSerHisThrSerProTyrPheProAsnArg 24
Db 1263 AATAAATCAAAATGAATGAAATATATAGATCGCACACTTCT-----ACTTCT 1310

QY 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
Db 1311 GTATCCAAATGATCTCAACATACATCCCTTTTTCGGAATGAGCCAAACAAATGCGCTACAAAAT 1370

QY 44 ThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
Db 1371 ATCGATTATAAAGATATTTAAATGCTCGCGGAATGCTAGTGAATACCCCTGGTTCA 1430

QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSer 82
Db 1431 CCTGAAGTACTGTTAGCGGACAAAGAT-----GCAGCTAAGCGCCCAATTGATATAGTA 1484

QY 83 GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIle 102
Db 1485 GGTAAATTTACTACAGGTTTAGG-----GTCCCAATTTGTT 1520

QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGlyGluGln 122
Db 1521 GGCCCGATAGTAGTCTTTTATCTCAACTATTGATATTCTGTGGCCTTCAGGCGCA-- 1577

QY 123 AspLysThrValTyrThrGlnPheIleLysMetGlyGlyIlePheValAspThrProLeu 142
Db 1578 ---AAGAGTCAATGGGAGATTTTATGGAACAAGTAGAAGACTCAATAATCAAAAATA 1634

QY 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
Db 1635 GCAGAAATATCAAGGAATAAAGCGCTTCGGAATTAGAGATTAGGTAATAATTACCAA 1694

QY 163 SerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
Db 1695 TTATATCTCACTCGCTTGAAGATGGAAAGAA----- 1727

QY 183 ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal 202
Db 1728 AATCCAAATGGTTCA-----AGAGCCTTAGAGATGTGGAATCCGATTGAAATCCCTG 1781

QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
Db 1782 GATAGTTATTATAGCAATACATGCCATCTTTTCGAGTGCACAAATTTTGAAGTACCATTC 1841

QY 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlyAlaGlu 242
Db 1842 CTTACAGTATATACACAGGCGACCACTTCATTACTGTTATTAAAGGAGCGCTTCAATT 1901

QY 243 LeuAlaAspGluTyrProAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
Db 1902 TTTGGAGAAGAATGGGGATG-----TCTACAAACCAT 1934

QY 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
Db 1935 ATTAATAACTATTATGATCGTCAATGAACCTTACTGAGAAATATCTGATCAGTCTGTGTA 1994

QY 283 AsnThrTyrArgThrGlyLeuLysAsnLeuAspGluProAsnMetLysTrpSerIle 302
Db 1995 AAGTGGTATGAACCTGTTTAGCAAAATTAAGAGCGCAGCGCTAAACAATGGGTGCGAC 2054

QY 303 PheAsnAspTyrArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db 2055 TATAACCAATTCCTAGAGAAATGACATGACGGTTTGTAGATGTTGTGCTATTTCCCA 2114

QY 323 LeuTyrAspIleLysArgTyrArgSerIleGlyGlyIleGluValLysGlyIleLys 342
Db 2115 AATTATGACACAGCAGCTAC-----CCAATGGAACAGAAACA 2153

QY 343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362
Db 2154 --CAACTAACAGGAAGATATATACA-----GATCCACTGGCGCGGTA 2195

QY 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
Db 2196 AACGTGCTCTCAATTTGGTTCTGCTATGACAAAGCACCTTCTTTCCGGAGTAGAATCA 2255

QY 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
Db 2256 TCGGTATTTCGACCAACC-----CATGTATTGATTATATAACGGGACTCACGTGATACA 2312

QY 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408
Db 2313 CAATCAAGACGATTTCTTCGCTCGCTATATAAGACATTTGGCTGGTTCATCAATAAGC 2372

QY 409 AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySer 428
Db 2373 TATCATCGGATTTTGTAGTAATATATAAAACAGATGTATGGAACCTAAATCAAAATCTA 2432

QY 429 ProThrThrLysThrIleArgProPheGluSerTyr-----Lys 441
Db 2433 CACAGCACTAGTACCTTT---GATTTTACGAATTATGATATTTTACAAGACGTTATCAAAA 2489

QY 442 ValSerIleValThrAspArgGlnSerProValSerProIleGlnProHisPheIle 461
Db 2490 GATCGGTGCTCTGTGATATTTGTTTTTCTGGTTATACGTATATA-----TTTTTGA 2543

QY 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
Db 2544 ATGCCAGAGTCGAGTTTTTCATGGTAAACCAATTTGAATATATACCAGAAAGACGTTAAAG 2603

QY 479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr 491
Db 2604 TATAATCCGGTTTCCAAAGATATATATAGCGGG----- 2636

QY 492 ThrPhePheGlnPheProArgLysAspCysAsnLeuValIleAspProGlyCysSer 511
Db 2637 -----ACAAGATTCGGAATTAGAAATACCTCCAGAAACTTCA 2675

QY 512 -----ProAsnPheAsnTyrSerHisIleLeuSerHisPhe---SerLeuPheThr 528
Db 2676 GATCAACCAATATGAGTCATATAGCATAGATTATGTCATATCACAAGTATTTCCCGCG 2735

QY 529 TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTrpThr 548
Db 2736 ACGGTTTCAACTACCGGATTA-----GTACCTGATTTTCTTGGACA 2777

QY 549 HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla 568
Db 2778 CATCGGATGCGGATCTTTATAAATGCAGTTCTATTCAGATAAAATTAATCAAGTATCCG 2837

QY 569 IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu 582
```

```

Db      2838 GTAAAGGTTTCTGATTTGGCTCCTCTATAAACAGGAGGCGCAATAATATACCGTTGTATCG 2897
Qy      583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGly----- 598
Db      2898 GGTCTGGATTTACAGGGGGGGGATAATAAAGTAATAAGAAATGAGTGAATATATATCA 2957
Qy      599 -----ArgLeuGluLeuThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616
Db      2958 CATATCGTGTAAAAATTTCA-----GACATTAAACAAAGAAATATAGTATGAGG 3005
Qy      617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
Db      3006 ATTCGGTATGCTTCC-----GCTAATAATACCTGAAATTTTATATAATCCTTCTGAAGAA 3059
Qy      637 GlyValIleGlyLeuProGlnArgLeuAsnThrPheSerGlyThrAsnTyrAsn 656
Db      3060 AACGTTAAATCT---CACGCTCAAAAACTAGTAAGAGGTGAAGCTTTAAACATATAAT 3116
Qy      657 AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValThr 673
Db      3117 AAATTTAAATATGCGACTTTGCCCTTATTAATTTTACGACAAACCGAACCTTTCAATTACT 3176
Qy      674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaIleValSerAsnSerIle 693
Db      3177 CTAGGGGCT-----ATATTTGAAGCGGAAGACTTTCTTGGAAATTGAA 3218
Qy      694 LeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGlu 713
Db      3219 GCTTATATAGCGAATCGAATTTATCCAGTAGATGAGACATAT-----GAAGCG 3269
Qy      714 LysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsn 730
Db      3270 GAACAGATTAGAACGACGCGAGAGAGCAAGTGAATCCCTTGTGTAGCAAT 3320

```

RESULT 15

```

US-10-606-320-18
; Sequence 18, Application US/10606320
; Publication No. US20040091505A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE REFERENCE: 35718/263948
; CURRENT APPLICATION NUMBER: US/10/606,320
; CURRENT FILING DATE: 2003-06-25
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 6613
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Genomic Cry12Ib-2
US-10-606-320-18

```

```

Alignment Scores:      1.47e-63      Length:      6613
Pred. No.:            703.00      Matches:      219
Score:

```

```

Percent Similarity: 44.53%      Conservative: 127
Best Local Similarity: 28.19%      Mismatches: 289
Query Match: 18.17%      Indels: 142
DB: 7      Gaps: 32

US-10-783-417-2 (1-735) x US-10-606-320-18 (1-6613)

Qy      5 AsnAspAsnAsnGlnTyrGluIleIleAspSerHisThrSerProTyrPheProAsnArg 24
Db      1263 AATAAATCAAAATGAATATGAATTTATAGATCGACACCTTCT-----ACTTCT 1310
Qy      25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
Db      1311 GTATCCAATGATTTCTAACAGATACCTTTTTCGGAATGAGCCCAACAAATTCGCTACAAAAT 1370
Qy      44 ThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
Db      1371 ATGATTTATAAGATTTATTTAAANAATGTCTCGGGGAATGCTAGTAATACCTCGGTCA 1430
Qy      63 PheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIleValSer 82
Db      1431 CCTGAAGTACTTGTTCGCGGCAAGAT-----GCAGCTAAGCGCGCAATTTGATATAGTA 1484
Qy      83 GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
Db      1485 GGTAAATTTACTATCAGCTTTTAGG-----GTCCCATTTTGT 1520
Qy      103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGlyGluGln 122
Db      1521 GGGCGATAGTAGTCTTTTACTCAACTTATTTGATTTCTGTGGCCTTTACGGGCA--- 1577
Qy      123 AspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu 142
Db      1578 ---AAGAGTCAATGGGAGATTTTATGGAACAAGTAGAAGAACTCATATAATCAAAAATA 1634
Qy      143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
Db      1635 GCAGAAATATCAAGGAATAAAGCGCTTTTCGAAATTAGAGGATTTAGGTAATAAATATACCAA 1694
Qy      163 SerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
Db      1695 TTATATCTAACTCGCTTGAAGAAATGGAAGAA----- 1727
Qy      183 ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal 202
Db      1728 AATCCAATATGTTCA-----AGAGCCTTAGAGATGTGGAATCGATTTGAAATCCTG 1781
Qy      203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysLeuLeu 222
Db      1782 GATAGTTTATTTACGCAATACATGCCATCTTTTCGAGTGACAAAATTTTGAAGTACCATTC 1841
Qy      223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlyAlaGlu 242
Db      1842 CTTACAGTATATACACAGGCGAGCAACCTTCATTACTGTATTATAAGAGGCGCTCAATT 1901
Qy      243 LeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
Db      1902 TTTGGAGAAATGGGGATCG-----TCTACACCACT 1934
Qy      263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
Db      1935 ATTAATAACTATTATGATCGTCAAACTGAACTTACTGCAGAAATATCTGATCCTGTGTA 1994
Qy      283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle 302
Db      1995 AAGTGGTATGAAACTGCTTTAGCAAAAATTAAGAGCGACGAGCGCTAAACAAATGGTTCGAC 2054
Qy      303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db      2055 TATAACCAATTCGCTAGAGAAATGACACTGACGGTTTTAGATGTTGTTCATTTATCCCA 2114
Qy      323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyIleGluValLysGlyIleLys 342

```

```
Db 2115 AATTATGACACGACGCTAC-----CCAAATGGAACCGAAAGCA----- 2153
Qy 343 AsnGluLeuThrArgGluLeuThrThrThrGluLeuPheAspArgLeuProGlnLeu 362
Db 2154 ---CAACTAAAGGAAGTATATACA-----GATCCACTGGGCGCGTA 2195
Qy 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
Db 2196 AAGTGTCCTCAATTGGTCTCGTATGACAAAGCCTCTTTCGGAGTGATGAGATCA 2255
Qy 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
Db 2256 TCGGTATTTCGACCAACC---CATGTATTGATTATATAACGGGACTCACAGTGTATACA 2312
Qy 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408
Db 2313 CAATCAAGAAGCATTTCTCGCTCGCTATATAAGACATTTGGGTGGTCAATAAAGC 2372
Qy 409 AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySer 428
Db 2373 TATCATCGGATTTTGTAGTATAATTATATAAACAAGATGTATGAACCTAAATCAAAATCTA 2432
Qy 429 ProThrThrLysThrIleArgProPheGluSerTyr-----Lys 441
Db 2433 CACAGCACTAGTACCTTT---GATTTTACGAAATTATGATATTTACAAGACGTTATCAAAA 2489
Qy 442 ValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIle 461
Db 2490 GATCGGTGTCCTTGATATTGTTTCTCGTTATACGTATATA-----TTTTTGGGA 2543
Qy 462 IleAsnGlnIleGluLeuThrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
Db 2544 ATCCAGAGACTCGAGTTTTCATGCTGAACCAATTTGAATAATACCAAGAACGCTTAAAG 2603
Qy 479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr 491
Db 2604 TATAATCCGGTTTCCAAAGATATATATAGCGGG----- 2636
Qy 492 ThrPhePheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSer 511
Db 2637 -----ACAAGAGATTCGGAATTAGAAATTACCTCCAGAAACTTCA 2675
Qy 512 -----ProAsnPheAsnTyrSerHisIleLeuSerHisPhe-----SerLeuPheThr 528
Db 2676 GATCAACCAAAATTATGAGTCATATAGCCATAGATTATGTCATATACAAAGTATTTCCGCGC 2735
Qy 529 TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThr 548
Db 2736 ACGGGTTCAACTACCGGATTA-----GTACCTGTATTTTCTTGGACA 2777
Qy 549 HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla 568
Db 2778 CATCGGAGTCCGATCTTATAATGCGAGTTCATTCAGATAAAATTACTCAGATTCGCGTC 2837
Qy 569 IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu 582
Db 2838 GTAAGGGTTCTGATTTGGTCTCCTCTAACAGAGCGGCCAAATAATACCGTTGTATCG 2897
Qy 583 GlyProGlyHisThrGlyLysLeuValTyrLeuGlnSerGlnGly----- 598
Db 2898 GGTCTCGATTTACAGGGGGGGGATATAAAGTAATAAGAAATGGAGTAATATATATCA 2957
Qy 599 -----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616
Db 2958 CATATCGTGTTTAAATTTCA-----GACATTAAACAAAGAAATATAGTAGG 3005
Qy 617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
Db 3006 ATTCGGTATGCTTCC-----GCTAATAAATACTGAATTTTATATAATCTCTCTGAAGAA 3059
Qy 637 GlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn 656
Db 3060 AACGTTAAATCT---CACGCTCAAAAAAATATGAATAGAGGTGAAGCTTTTAAACATATAAT 3116
```

```
Qy 657 AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValThr 673
Db 3117 AAATTTAATTATGCGACTTTGCCCCCTATTAAATTTACGACACCGAACCTTTTCATTACT 3176
Qy 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693
Db 3177 CTAGGGGCT-----ATATTGAAGCGGAAGACTTTCTTGGAAATTGAA 3218
Qy 694 LeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGlu 713
Db 3219 GCTTATATAGACCGGAATCGAATTTTATCCAGTATGATGAGACATAT-----GAAGCG 3269
Qy 714 LysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsn 730
Db 3270 GAACAAGATTTAGAAAGCAGCGAAGAAAGCAGTGAATGCCTTGTGTACGAAT 3320
```

Search completed: December 11, 2005, 23:17:11
Job time : 1528 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 11, 2005, 20:06:37 ; Search time 173 Seconds

(without alignments)
1588.342 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869

Sequence: 1 MNQNDNNEYIIDSHTSPY.....KLETIQTKINTPTNHTKTL 735

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3392430 seqs, 186927314 residues

Total number of hits satisfying chosen parameters: 6784860

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-O=/cgn2_1/USPTO_spool/US10783417/runat_01122005_141955_24269/app_query.fasta_1.903
-DB=Published Applications NA.New -QWMT=fastap -SUFFIX=rnpbn -MINMATCH=0.1
-LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10783417 @CGN 1 139 @runat_01122005_141955_24269
-NCPUS=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOF=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- Published Applications NA.New:*
1: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	703	18.2	3633	US-11-058-727-3	Sequence 3, Appli
2	703	18.2	3633	US-11-108-389-3	Sequence 3, Appli
3	703	18.2	6613	US-11-058-727-18	Sequence 18, Appl
4	703	18.2	6613	US-11-108-389-18	Sequence 18, Appl
5	689.5	17.8	4188	US-11-091-643-5	Sequence 5, Appli
6	678.5	17.5	2022	US-11-058-727-13	Sequence 13, Appl
7	678.5	17.5	2022	US-11-108-389-13	Sequence 13, Appl
8	675	17.4	3621	US-11-058-727-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-11-058-727-3
; Sequence 3, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:

- ; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal

; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3

; LENGTH: 3633

; TYPE: DNA

; ORGANISM: Bacillus thuringiensis

; FEATURE:

Sequence 1, Appli
Sequence 17, Appl
Sequence 3, Appli
Sequence 73, Appl
Sequence 79, Appl
Sequence 73, Appl
Sequence 79, Appl
Sequence 79, Appl
Sequence 81, Appl
Sequence 81, Appl
Sequence 41, Appl
Sequence 41, Appl
Sequence 47, Appl
Sequence 47, Appl
Sequence 49, Appl
Sequence 49, Appl
Sequence 69, Appl
Sequence 69, Appl
Sequence 33, Appl
Sequence 67, Appl
Sequence 33, Appl
Sequence 67, Appl
Sequence 75, Appl
Sequence 75, Appl
Sequence 21, Appl
Sequence 63, Appl
Sequence 21, Appl
Sequence 63, Appl
Sequence 63, Appl
Sequence 65, Appl
Sequence 77, Appl
Sequence 77, Appl
Sequence 43, Appl
Sequence 43, Appl
Sequence 7, Appl
Sequence 7, Appl
Sequence 45, Appl

```

; NAME/KEY: CDS
; LOCATION: (1)...(3633)
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Cry1218-2
US-11-058-727-3

```

Alignment Scores:	
Pred. No.:	2,09e-64
Score:	703.00
Percent Similarity:	44.53%
Best Local Similarity:	28.19%
Query Match:	18.17%
DB:	7
	Gaps:
	32
	Indels:
	142
	Mismatches:
	289
	Conservative:
	127
	Matches:
	219
	Length:
	3633

US-10-783-417-2 (1-735) x US-11-058-727-3 (1-3633)

QY	5	AsnAspAsnGlnTyrGluIlelleAspSerHisThrSerProTyrPheProAsnArg	24
DB	10	AATAAATCAAAATGAATATGAATATATAGATGCACACTCTCT-----ACTTCT	57
QY	25	AsnSerAsnAspSer--AtqTyvProTyrThrAsnAsnProAsnGlnProLeuGlnAsn	43
DB	58	GTAATCCAAATGATTCTAACAGATACCCTTTTCGGAATGAGCCCAAAATGCGCTACAAAAT	117
QY	44	ThrAsnTyrLysGluTrpLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn	62
DB	118	ATGGATTATAAAGATTATTTAAAAATGTCTCGGGAAATGCTAGTAGAATACCTCGTGTCCA	177
QY	63	PheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIleValSer	82
DB	178	CCTGAAGTACTTGTTAGCGGACAAGAT-----GCAGCTAAGCGCGCAATTGATATAGTA	231
QY	83	GlyThrLeuLeuAlaGlyIleGlyCylLeuThrSerIleSerGlyProIleGlyIleIle	102
DB	232	GGTAATLTACTATCAGGTTTAGG-----GTCCCATTTGTT	267
QY	103	GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGln	122
DB	268	GGGCCGATAGTAGCTCTTTACTCAACTTATTGATATTCGTGGCCTTCAGGGCAA---	324
QY	123	AspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu	142
DB	325	---AAGAGTCAATGGGAGATTTTTATGGAAACAAGTAGAAGAAGCTCATATAATCAAAAAAT	381
QY	143	ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln	162
DB	382	GCAGAAATGCAAGGAATAAGCGCTTTCGGNAATTAGAAGGATTAGGTAAATAATTACCAA	441
QY	163	SerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeu	182
DB	442	TTATATCTAATCGCGCTTGAAGAATGGAAGAA-----	474
QY	183	ProProSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal	202
DB	475	AATCCAATGGTTCA-----AGAGCCCTTACAGATGTCGGAATCGATTTGAAATCCCG	528
QY	203	HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu	222
DB	529	GATAGTTTATTACGCAATACATGCCATCTTTTCGAGTGACAAAATTTGAAGTACCATTTC	588
QY	223	LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGlu	242
DB	589	CTTACAGTATATACAGCGCACCAACCTCATTTACTTGTTATTAAAGACGCTTCAAPT	648
QY	243	LeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr	262
DB	649	TTTGGAGAGNATGGGATGG-----TCTACAACCACT	681
QY	263	SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla	282
DB	682	ATTAATAACTATTATGATCGTCAATAGAACTTACTGCAGATATCTTGATCACTGTGTA	741

283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle 302
Db AAGTGGTATGAACCTGCTTTAGCAAAATTAAAGACGACGAGCGCTAAACAATCGGTGCAC 801
303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db TATAACCAATTCCTCGTAGAAGAAATGACACTGACCGTTTATAGATGTGTGTGATATTATCCCA 861
323 LeuTyrAspIleLysArgTyrArgSerIleGlyGlyIleGluValLysGlyIleLys 342
Db AATTATGACACACGCGTAC-----CCATGGAAACGAAGCA----- 900
343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362
Db ---CAACTAAACAGGAAGTATATACA-----GATCCACTGGCGCGGTA 942
363 ArgValGln-----ProAsnLeuIleAlaThrMetGluTyr 373
Db AAGCTGTCTTCAATGGTTCTCTGCTATGACAAAGCACCTTCTTTCGGAGTGATAGATCA 1002
374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
Db TCGGTATTCGACACCC---CATGTATTGTATTATTAACGGGACTCAGGTGTATACA 1059
394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408
Db CAATCAAGAGCATTTCTTCGCTCGCTATATAAGACATTTGGCGTGGTCATCAATAAGC 1119
409 AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySer 428
Db TATCATCGGATTTTATAGTGAATAATATTATAAAACAGATGATGTAACCAATAATCAAAATCTA 1179
429 ProThrThrLysThrIleArgProPheGluSerTyr-----Lys 441
Db CACAGCACTAGTACCTTT---GAATTTACGAATTTATGATATTTACAGACGTTATCAAAA 1236
442 ValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIle 461
Db GATCGGGTCTCTGATATTGTTTTCTGGTTATACGTATATA-----TTTTTTGGA 1290
462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
Db ATGCCAAGACTCGAGTTTTTTCATGGTAAACCAATTCGAATAATACCAGAAAGACGTTAAAG 1350
479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr 491
Db TATAATCCGGTTTCCAAAGATATATTAGCGGGG----- 1383
492 ThrPhePheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSer 511
Db ---ACAAGAGATTCGAATTTAGAAATTACTCTCAGAAACTTCA 1422
512 -----ProAsnPheAsnAntyTrpSerHisIleLeuSerHisPhe---SerLeuPheThr 528
Db GATCAACCAAAATTATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCGCGG 1482
529 TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThr 548
Db ACGGGTTCAACTACCGGATTA-----GTACCTGTATTTTCTTCGACA 1524
549 HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla 568
Db CATCGGAGTCCCGATCTTTATAAATGCAAGTTTCATTCAGATAAAATTAATCTCAGTTCGGT 1584
569 IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu 582
Db GTAAAGTTTCTGATTTGGCTCCCTCTATAACAGGAGGCCAAATAATACCGTTGTATCG 1644
583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGly----- 598
Db GGTCTCTGATTTACAGGGGGGGGATATAAAAGTAATAAGAAATGAGTAATATATATCA 1704


```

QY 599 -----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616
Db 1705 CATATCGGTGTTAAATTTCA-----GACATTAAACAAGATATAGTATGAGG 1752
QY 617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
Db 1753 ATTCGGTATGCTTCC-----GCTAATAATACCTGAAATTTTATATAAATCCTTCTCTGAAGAA 1806
QY 637 GlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn 656
Db 1807 AAGCTTAATATCT---CACGCTCAAAAACCTATGAATAGAGGTGAAGCTTTTAAACATATAAT 1863
QY 657 AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValThr 673
Db 1864 AAATTTAATATCGACTTGGCCCTTATTAATTTACGACACCGCAACCTTTCAATTAAT 1923
QY 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693
Db 1924 CTAGGGGCT-----ATATTTGAAGCGGAGACATTCTTCTGGAATTTGAA 1965
QY 694 LeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGlu 713
Db 1966 GCTTATATAGCGAATCGAATTTATCCAGTAGATGAGACATAT-----GAAGCG 2016
QY 714 LysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsn 730
Db 2017 GAACAAGATTTAGACGCGAAGAAAGCAGTGAATGCTTGTGTTACGAAT 2067

```

RESULT 2

```

US-11-108-389-3
; Sequence 3, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F. H. Wong
; APPLICANT: Cao-Quo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 357118/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3633
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3633)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Cry1218-2
US-11-108-389-3

```

```

Alignment Scores:
Pred. No.: 2,09e-64 Length: 3633
Score: 703.00 Matches: 219
Percent Similarity: 44.53% Conservativity: 127
Best Local Similarity: 28.19% Mismatches: 289

```

```

Query Match: 18.17% Indels: 142
DB: 7 Gaps: 32
US-10-783-417-2 (1-735) x US-11-108-389-3 (1-3633)
QY 5 AsnAspAsnAsnGluTyrGluIleAspSerHisThrSerProTyrPheProAsnArg 24
Db 10 AATAATCAAAATGAATATGAATATATAGATCGCACCTTCT-----ACTTCT 57
QY 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
Db 58 GTATCCAATGATTTCTAAACAGATACCTTTTTCGAATAGAGCCAAACAAATCGGTACAAAAT 117
QY 44 ThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
Db 118 ATGATTATTAAGATTATTTAAATAATGCTCGCGGAATGCTAGTAATACCTCGGTTCAT 177
QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIleValSer 82
Db 178 CCTGAAGTACTTGTAGCGGCAAGAT-----GCAGCTAAGCGCGCAATTGATATAGTA 231
QY 83 GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
Db 232 GGTAAATTAATCTATCAGCTTTAGGG-----GTCCCATTTGTT 267
QY 103 GlyAlaIleIleIleSerPheGlyThrIleIleThrValPheThrProAlaGlyGluGln 122
Db 268 GGGCCGATAGTAGTCTTTTATCTCAACTTATTGATATTCTGTGGCTTCAGGGCAA--- 324
QY 123 AspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu 142
Db 325 ---AAGAGTCAATGGGAGATTTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATA 381
QY 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
Db 382 GCAGATATCAAGGAATAAAGCCCTTTTCGAATTAGAAGGATTAGGTAAATAATTACCAA 441
QY 163 SerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
Db 442 TTATATCTAATCTGCGCTTGAAGATGGAAGAA----- 474
QY 183 ProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGluAsnVal 202
Db 475 AATCCAAATGGTTCA-----AGAGCCCTTAGCAGATGTGCGAAATCGAATTTGAAATCCTG 528
QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
Db 529 GATAGTTTATTAGCGCAATACATCCCATCTTTTCGAGTGACAAATTTTGAAGTACCATTC 588
QY 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGlu 242
Db 589 CTTACAGTATATACACAGCGAGCCAACTTCATTCTGTTTATTAAAGGACGCTTCAATT 648
QY 243 LeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
Db 649 TTTGGAGAAGAATGGGGATGG-----TCTACAACCACT 681
QY 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
Db 682 ATTAATAACTATTATGATCGTCAAAATGAACCTTACTGCAGAAATATTCTGATCCTGTGTA 741
QY 283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle 302
Db 742 AAGTGGTATGAAACTGTTTGTAGCAAAATTAAGGCGCAGCGCTTAAACATGGGTGCGAC 801
QY 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db 802 TATAACCAATTCGCTAGAGAAATGACACTGACGCTTTTAGATGTTGTGCTATTATCCCA 861
QY 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyIleGluValLysGlyIleLys 342
Db 862 AATTATGACACACGACGCTAC-----CCAATGGAAACGAAAGCA----- 900

```

```
QY 343 AsnGluLeuThrArgGluLeuThrThrGluLeuAsnPheAspArgLeuProGlnLeu 362
Db 901 ---CAACTAACAGGGAAGTATATACA-----GATCCACTGGCGCGGTA 942
QY 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
Db 943 AAGGTGCTTCAATGGTTCCTGGTATGACAAAGCACCTTCTTCGGAGTGATAGATCA 1002
QY 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheLeuThrThr 393
Db 1003 TCCGTTATTTCGACCAACC---CATGTATTGTATTATATAACAGGGACTCCAGGTGTATACA 1059
QY 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyLysSerAsnArg 408
Db 1060 CAATCAAGAAGCATTTCTTCGCTCGCTATATAGAATGGCTGGCTCATCAATAAGC 1119
QY 409 AspAlaProThrTyrSerAsnThrThrLeuThrLeuTyrGlyGluArgThrGlySer 428
Db 1120 TATCATCGGATTTTAGTGATGATATATATAAACAGATGATGGAACATAATCAAAATCTA 1179
QY 429 ProThrThrLysThrLysArgProPheGluSerTyr-----Lys 441
Db 1180 CACAGCACTAGTACCTTT---GATTTACGAATTATGATATTTACAAGAGCTTATCAAA 1236
QY 442 ValSerIleValThrAspArgGlnSerProValSerProIleGlnProHisPheIle 461
Db 1237 GATCGGTGCTCCTTGATATGTTTTCTCTGGTTATACGTATATA-----TTTTTGG 1290
QY 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
Db 1291 ATGCCAGAGTCGAGTGTTCATGGTAAACCAATTGAATAATACCAGAAAGAGCTTAAAG 1350
QY 479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr 491
Db 1351 TATAATCCGGTTTCCAAAGATATATAGCGGG-----1383
QY 492 ThrPhePheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSer 511
Db 1384 -----ACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCA 1422
QY 512 -----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPhe---SerLeuPheThr 528
Db 1423 GATCAACCAAAATTATGAGTCATATAGCCATAGATTATGTATATACCAAGATATCCCGG 1482
QY 529 TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThr 548
Db 1483 ACGGGTTCMACTACCGGATTA-----GTACCTGTATTTCTTGGACA 1524
QY 549 HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla 568
Db 1525 CATCGGAGTCCGATCTTATAATGCGAGTTCATTCAGATAAAATTAATCAGATTCGGTC 1584
QY 569 IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu 592
Db 1585 GTAAAGGTTTCTGATTTGGCTCCCTCTATAACAGGAGGCCAAATAATACCGTTGATCG 1644
QY 593 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGly-----598
Db 1645 GGTCTCGGATTTACAGGGGGGGGATATAAAGTAATAAGAAATCGGATTAATATATCA 1704
QY 599 -----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616
Db 1705 CATATCGGTGTTAAATTTCA-----GACATTAACAAAGAAATATAGTAGGAGG 1752
QY 617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
Db 1753 ATTCGGTATGCTTCC-----GCTAATAATCTGAAATTTTATATAAATCCTCTCGAAGAA 1806
QY 637 GlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn 656
Db 1807 AACGTTAAATCT---CACGCTCAAAAACTATGAATAGAGTGAAGCTTTAAACATATAAT 1863
QY 657 AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValThr 673
```

```
Db 1864 AAATTTAATTATCGCACTTTGCCCCCTATTAAATTTAGCAACACCGAACCTTTCTACT 1923
QY 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693
Db 1924 CTAGGGGCT-----ATATTTGAAGCGGAGAGACTTTCTTGGATTTGAA 1965
QY 694 LeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGlu 713
Db 1966 GCTTATATAGACCGAATCGAATTTATCCAGTAGATGAGACATAT-----GAAGCG 2016
QY 714 LysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsn 730
Db 2017 GAACAAGATTTAGACGAGCAGCAAGAAAGACAGTGAATGCTTGTTTACGAAT 2067

RESULT 3
US-11-058-727-18
; Sequence 18, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; PRIOR FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 6613
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Genomic Cry12Ib-2
US-11-058-727-18

Alignment Scores:
Pred. No.: 5,54e-64 Length: 6613
Score: 703.00 Matches: 219
Percent Similarity: 44.53% Conservative: 127
Best Local Similarity: 28.19% Mismatches: 289
Query Match: 18.17% Indels: 142
DB: 7 Gaps: 32

US-10-783-417-2 (1-735) x US-11-058-727-18 (1-6613)
QY 5 AsnAspAsnAsnGluTyrGluIleAspSerHisThrSerProTyrPheProAsnArg 24
Db 1263 AATAATCAAAATGAATATGAATATATAGATGCGACACCTTCT-----ACTTCT 1310
QY 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
Db 1311 GTATCAATGATTTCTAATACAGATACCTTTTGGCAATGAGCAACAATCGCTACAAAT 1370
QY 44 ThrAsnTyrLysGluTrpLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
Db 1371 ATGGATTATAAAGATTATTTAAAAATGCTCGCGGAAATGCTAGTAGAATACCTCGTTTCA 1430
```

QY 63 PheGluThrPheAlaSerAlaAAspThrIleAlaValSerAlaGlyThrIleValSer 82
Db 1431 CCGAAGTACTTGTACGGACAAAGAT-----GCAGTAAAGCCGCAATGATGATAGTA 1484
QY 83 GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIle 102
Db 1485 GGTAAATTAATCACTACAGGTTTAGG-----GTCCCAATTGTT 1520
QY 103 GlyAlaIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGln 122
Db 1521 GGGCCGATAGTACTTGTATCTCAACTGATATCTATTGATATCTGTGGCTTCAGGGCAA-- 1577
QY 123 AspLysThrValThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu 142
Db 1578 ---AAGAGTCAATGGGAGATTTTATGGAAACAAGTAGAAGCACTCAATAATCAAAAATA 1634
QY 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
Db 1635 GCAGATATGCAAGGAATAAGCGCTTTCGGAATTAGAGGATTAGGTAATATTACCAA 1694
QY 163 SerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
Db 1695 TTATATCTAATCGCTTGAAGATGAAAGAA----- 1727
QY 183 ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal 202
Db 1728 AATCCAAATGGTTCA-----AGAGCCCTTACGAGATGTGCGAATCGATTGGAATCCTG 1781
QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
Db 1782 GATAGTTATTAGCAATACATCGCCATCTTTCGAGTGACAAATTTGAAGTACCATTC 1841
QY 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGlyAlaGlu 242
Db 1842 CTTACAGTATATACACAGCGAGCACTTCAITTTACTGTTTAAAGACGCTTCAATT 1901
QY 243 LeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
Db 1902 TTTGGAGAGAAGATGGGGATGG-----TCTACAAACCACT 1934
QY 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
Db 1935 ATTAATAACTATTATGATCGTCAAAATGAACTTACTGCGAATAATCTGATCATCTGTGA 1994
QY 283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle 302
Db 1995 AAGTGGTATCAACTGGTTTAGCAAAATTAAGGACGAGCGCTAAACAATGGGTGCAC 2054
QY 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db 2055 TATAACCAATTCGTTAGAGAAATGACACTGACGGTTTATAGATGTTGTCATTATTCCTCA 2114
QY 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIleLys 342
Db 2115 AATTATGACACACGACGCTAC-----CCAAATGGAAACGAAGA----- 2153
QY 343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362
Db 2154 ---CAACTAACAGGAGAGTATATACA-----GATCCACTGGCGCGGTA 2195
QY 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
Db 2196 AAGCTGTCTCAATTGCTTCTGTTATGACAAAGCACCTTCTTTCGGAGTGATAGATCA 2255
QY 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
Db 2256 TCCGTTATTCCGACACCC---CATGTATTTGATTATATATAACGGGACTCACAGTGTATACA 2312
QY 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408
Db 2313 CAATCAAGAGACATTTCTTCCGTCGCTATATAAGACATTTGGGCTGCTCATCAATAAGC 2372

RESULT 4

QY 409 AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySer 428
Db 2373 TATCATCGGATTTTAGTGATAATATTATAAACACAGATGTATGGAACATAATCAAAATCTA 2432
QY 429 ProThrThrLysThrIleArgProPheGluSerTyr-----Lys 441
Db 2433 CACAGCACTAGTACTT---GATTTTACGAATATGATATTATTACAGACGTTATCAAAA 2489
QY 442 ValSerIleValThrAspArgGlnSerProProValGlnSerProIleGlnProHisPheIle 461
Db 2490 GATCGGCTGCTCTGATATATTGTTTTCTGTTATACGTATATA-----TTTTTTGGA 2543
QY 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
Db 2544 ATGCCAGAAGTCGAGTTTTTCATGTTAAACCAATTTGAATAATACCCAGAAAGACGTTAAAG 2603
QY 479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr 491
Db 2604 TATAATCCGGTTTCCAAAGATATTATAGCGGG----- 2636
QY 492 ThrPhePheGlnPheProArgLysAspCysAsnLeuValIleAspProGlyCysSer 511
Db 2637 -----ACAAGACATTCGGAATTAGAAATTACCTCCAGAAACTTCA 2675
QY 512 -----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPhe---SerLeuPheThr 528
Db 2676 GATCAACCAAAATTTAGTTCATATAGCATGATTTATGTCATATACAGATTAATCCCGCG 2735
QY 529 TyrSerTyrValIleGlyLeuGlnLeuAspThrGlyValLeuGlyTrpThr 548
Db 2736 ACGGTTCACCTACCGGATTA-----GTACCTGTATTTTCTTGGACA 2777
QY 549 HisSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla 568
Db 2778 CATCGAGTGCAGTCTTATAAATGCAGTTCATTACAGATAAAATTTACTCAGATTCGGTTC 2837
QY 569 IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu 582
Db 2838 GTAAAGGTTTCTGATTTGGCTCCCTCTATAACAGAGGCGCAATAATAATACCTTGTATCG 2897
QY 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGly----- 598
Db 2898 GGTCTCGATTTACAGGGGGGGGATATAAAGTAATAAGAAATGGAATTAATATATATCA 2957
QY 599 -----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616
Db 2958 CATATGCTGTAAATTTCA-----GACATTACAAAGAATATAGTATGAGG 3005
QY 617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
Db 3006 ATTCGGTATGCTTCC-----GCTAATAATACTGAATTTTATATAAATCTTCTGAAGAA 3059
QY 637 GlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn 656
Db 3060 AAGCTTAAATCT---CACGCTCAAAAACCTATGAATAGAGTGAGGCTTTTAACATATAT 3116
QY 657 AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValThr 673
Db 3117 AAATTTAATTATGCGACTTTGCCCTTATTAATTTACGACAAACCGAACCTTTCATTACT 3176
QY 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693
Db 3177 CTAGGGCT-----ATAATTGAAGCGGAAGACTTTTCTTGGAAATTGAA 3218
QY 694 LeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGlu 713
Db 3219 GCTTATATAGACGAATCGAATTTTATCCAGTAGATGAGACATAT-----GAAAGC 3269
QY 714 LysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsn 730
Db 3270 GAACAGATTAGAACGACGAGAAAGACGTGAATGCTTGTTTTACGAAT 3320

```

US-11-108-389-18
; Sequence 18, Application US/11108389
; Publication NO. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 6613
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Genomic Cry1218-2
US-11-108-389-18

Alignment Scores:
Pred. No.: 5,54e-64 Length: 6613
Score: 703.00 Matches: 219
Percent Similarity: 44.53% Conservative: 127
Best Local Similarity: 28.13% Mismatches: 289
Query Match: 18.17% Indels: 142
DB: 7 Gaps: 32

US-10-783-417-2 (1-735) x US-11-108-389-18 (1-6613)
QY 5 AsnAspAsnGluTyrGluIleleAspSerHisThrSerProTyrPheProAsnArg 24
DB 1263 AATAATCAAAATGAATATGAAATATATAGATGCGACACCTTCT-----ACTTCT 1310
QY 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
DB 1311 GTATCAATGATCTTAACAGATACCTTTTGGGAATGAGCCAAATGCGGTACAAAAT 1370
QY 44 ThrAsnTyrLysGluTrpLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
DB 1371 ATGGATTATAAGATTTATTTAAATGTCTCGGGAATGCTAGTGAATACCTGTTCA 1430
QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSer 82
DB 1431 CCTGAAGTACTTGTAGCGGACAAGAT-----GCAGCTAAGGCGCAATTTGATATAGTA 1484
QY 83 GlyThrLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
DB 1485 GGTAATAATCACTACGATTTAGGG-----GTCCCAATTTGTT 1520
QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGln 122
DB 1521 GGGCGATAGTAGCTTTATCTACTCACTTTATGATTTCTGTGGCTTCAGGGGAAA--- 1577
QY 123 AspLysThrValTrpThrGlnPheIleLysMetGlyIlePheValAspThrProLeu 142
DB 1578 ---AAGAGTCAATGGGAGATTTTATGGAAACAAGTAGAAGAACTCATATAATCAAAAATA 1634

```

```

QY 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
DB 1635 GCAGAAATATCAAGGAATAAAGCGCTTCGGAATTAGAGGATTAGGTAAATAATACCAA 1694
QY 163 SerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
DB 1695 TTATATCTTAATCGCTTGAAGATGGAAGAA----- 1727
QY 183 ProProSerSerAlaLeuGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal 202
DB 1728 AATCCAAATGGTTCA-----AGAGCCTTAGGAGATGCGGAATCGATTTGAATCCTG 1781
QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
DB 1782 GATAGTTTATTACGCAATACATCCCATCTTCGAGTGACAAATTTGAAGTACCATTTC 1841
QY 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnAlaGlu 242
DB 1842 CTTACAGTATATACAGGCGACCACTTCATTTACTGTATTATAAGGACGCTTCAATT 1901
QY 243 LeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
DB 1902 TTGGAGAAAGATGGGGATGG-----TCTACAAACCACT 1934
QY 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
DB 1935 ATTAATAACTATTATGATCGTCAATGAACCTTACTCGAGAAATTTCTGATCAGTGT 1994
QY 283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle 302
DB 1995 AAGTGTATGAAACTGTTTTAGCAAAATTAAGGCGACGCGCTAAACAATGGTGCAC 2054
QY 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
DB 2055 TATAACCAATTCGTAGAGAAATGACACTGACGGTTTTAGATGTTGTTCATTTATCCCA 2114
QY 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyIleGluValLysGlyIleLys 342
DB 2115 AATTATGACACACGACGATAC-----CCAAATGGAACGMAAGCA----- 2153
QY 343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362
DB 2154 ---CAACTAACAGGGAAGTATATACA-----GATCCACTGGGCGCGTA 2195
QY 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
DB 2196 AAGCTGTCTTCAATGGTTCCTGGTATGCAAAAGCACCTTCTTTCGGAGTGATAGAATCA 2255
QY 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
DB 2256 TCGGTTATTCGACCAACC---CATGTATTGTATTATATACGGGACTCACAGTGTATACA 2312
QY 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408
DB 2313 CAATCAAGAAGCATTTCTTCGCTCGCTATATAAGACATTTGGGCTGTCATCAATAAGC 2372
QY 409 AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySer 428
DB 2373 TATCATCGGATTTTATGTATATATATATAAACAAGATGATATGGAACCTAATAATCTA 2432
QY 429 ProThrThrLysThrIleArgProPheGluSerTyr-----Lys 441
DB 2433 CACAGCACTAGTACCTTT---GATTTTACGAATTTATGATATTATACAGACGCTTATCAAA 2489
QY 442 ValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIle 461
DB 2490 GATGCGGTGCTCCTTGATATTTGTTTCTCTGTTTATACGTATATA-----TTTTTGA 2543
QY 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrIleLys 478
DB 2544 ATGCCAGAGTTCGAGTGTTCATGGTAAACCAATTTGAATAATACCAAGAACGCTTAAG 2603

```

```
QY 479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr 491
|||
|||
|||
Db 2604 TATAATCCGGTTTCCAAAGATATTATAGCGGG----- 2636
QY 492 ThrPhePheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSer 511
|||
|||
|||
Db 2637 -----ACAAGAGATTCGGAATTAGAAATACCTCCAGAAACTTCA 2675
QY 512 -----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPhe---SerLeuPheThr 528
|||
|||
|||
Db 2676 GATCAACCAAAATATAGTCATATAGCATATGTCATATATCAACAAGTATTCGCGG 2735
QY 529 TyrSerTyrValIleGlyLeuGlnLeuIleAspThrGlyValLeuGlyTyrThr 548
|||
|||
|||
Db 2736 ACGGTTTCAACTACCGAATTA-----GTACCTGATTTTCTTGACA 2777
QY 549 HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla 568
|||
|||
|||
Db 2778 CATCGGAGTCCGATCTTATAATGCGATTCATTCAGATAAAATTAATCAGATTCGCGTC 2837
QY 569 IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu 582
|||
|||
|||
Db 2838 GTAAAGGTTTCTGATTTGGCTCCCTCTATAAACAGGAGGCGCAATAATAACCGTTGTATCG 2897
QY 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGly----- 598
|||
|||
|||
Db 2898 GGTCTCGAATTTACAGGGGGGGGTAATAAAAGTAATAAGAAATGAGTAATTATATCA 2957
QY 599 -----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616
|||
|||
|||
Db 2958 CATATCGGTTTAAATTTCA-----GACATTAACAAGAAATATAGTATGAGG 3005
QY 617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
|||
|||
|||
Db 3006 ATTCGGTATGCTTCC-----GCTAATAATAGTAAATTTATATAATCCTTCTGAAGAA 3059
QY 637 GlyValIleGlyLeuProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn 656
|||
|||
|||
Db 3060 AAGCTTAAATCT---CACGCTCAAAAAAATCATGAATAGAGTGAAGCTTTAAACATATAAT 3116
QY 657 AsnLeuGluTyrClyAspPheGlyTyrPheGlnPheProSerThr-----ValThr 673
|||
|||
|||
Db 3117 AAATTTAATATGCGACTTGGCCCTATTAAATTTACGACACCGAACCCTTTCATTACT 3176
QY 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnAlaPheValSerAsnSerIle 693
|||
|||
|||
Db 3177 CTAGGGGCT-----ATATTTGAAGCGGAAGACTTCTTGGAAATGAA 3218
QY 694 LeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGlu 713
|||
|||
|||
Db 3219 GCTTATATAGCCGATCGAATTTATATCCAGTAGATGAGACATAT-----GAAGCG 3269
QY 714 LysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsn 730
|||
|||
|||
Db 3270 GAACAAGATTTAGAACGAGCGAAGAAAGCAGTGAATGCTTGTTCAGAAAT 3320
RESULT 5
US-11-091-643-5
; Sequence 5, Application US/11091643
; Publication No. US20050246789A1
; GENERAL INFORMATION:
; APPLICANT: TANAKA, Masao
; APPLICANT: YOKOYAMA, Tomoko
; APPLICANT: AOYAGI, Morichi
; APPLICANT: HASEGAWA, Makoto
; APPLICANT: EHARA, Gaku
; APPLICANT: KIMURA, Masaharu
; APPLICANT: NISHIHASHI, Hideji
; TITLE OF INVENTION: Polypeptide having larval growth inhibiting or
; TITLE OF INVENTION: insecticidal effect on scarabaeidae insects and
; TITLE OF INVENTION: polynucleotide encoding the same
; FILE REFERENCE: Op1335
; CURRENT APPLICATION NUMBER: US/11/091,643
```

```
; CURRENT FILING DATE: 2005-03-29
; PRIOR APPLICATION NUMBER: JP 2001-115754
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: JP 2001-203463
; PRIOR FILING DATE: 2001-07-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 4188
; TYPE: DNA
; ORGANISM: Bacillus popilliae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4158)
; US-11-091-643-5

Alignment Scores:
Pred. No.: 7.5e-63 Length: 4188
Score: 689.50 Matches: 239
Percent Similarity: 45.07% Conservative: 113
Best Local Similarity: 30.60% Mismatches: 286
Query Match: 17.82% Indels: 143
DB: 9 Gaps: 38

US-10-783-417-2 (1-735) x US-11-091-643-5 (1-4188)
QY 1 MetAsnGln-----AsnAsnAspAsn-----AsnGluTyr 10
|||
|||
|||
Db 55 ATGAATCAATATCATACCAAGAACGATAACAAAGTTACAAACCAAGTGAAGTGAAGTG 114
QY 11 GluIleIle-----AspSerHisThrSerProTyrPheProAsnArgAsnSerAsnAsp 28
|||
|||
|||
Db 115 CAAATTATACAAACCTTCAAGTAACGCTTTACTTTACAGTCCCAAC----- 159
QY 29 SerArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsnThrAsnTyrLysGlu 48
|||
|||
|||
Db 160 ---AAGTATCCGATGCTACGAGTCCCAATGTCATAGCAGAGGGTAGAAGTTATAAAAT 216
QY 49 TrpLeuAsnMetCysGlnGlyAsnThrGlnTyrGlyAspAsnPheGluThrPheAlaSer 68
|||
|||
|||
Db 217 TGGTTGATATGTGTAGGT-----GTAGGTGACGAT-----ACA 252
QY 69 AlaAspThrIleAlaAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAlaGly 88
|||
|||
|||
Db 253 CGAAGTCCCGAAGTCCGGTTACTGCACAAAGTTCTATTTCGACTTCTCTT----- 303
QY 89 IleGlyClyLeuThrSer---IleSerGlyProIleGlyIle-----IleGlyAlaIle 105
|||
|||
|||
Db 304 -----GGTATAACCAAGCAATCATTTGGCGCTCTAGGTATTCGCGTTGTAGCGGAAGCC 357
QY 106 IleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGlnAspLysThr 125
|||
|||
|||
Db 358 ATTGGAATTTTCGCGCAGCTATTGGATTGGTTGGCTCTAGGTATTCGCGTTGTAGCGGAAGCC 411
QY 126 ValTrpThrGlnPheIleLysMetGlyLysIlePheValAspThrProLeuThrGluSer 145
|||
|||
|||
Db 412 ---TGGGTATATTTATGAATCATGTAGAGAACTCATAAATTCAAAATAACAGAGACT 468
QY 146 IleLysGlnLeuLysLeuGlnThrLeuGluClyPheArgGlnIleLeuGlnSerTyrAsn 165
|||
|||
|||
Db 469 GTAAAAAATGAGGCAATTAACAAGATTAGACGGTTTATAGGCAATGTCTAGCGCTATATCAA 528
QY 166 ThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeuProProSer 185
|||
|||
|||
Db 529 AAGGCTTTCGAAGAATGCGAA-----CAACATCCT----- 558
QY 186 SerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAsp 205
|||
|||
|||
Db 559 ---ACTCTTGAATCAGCTCCAGTACGAGTGTACACATGATTTTTCTTAATGTAATAATTT 615
QY 206 PheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeuLeuProfile 225
|||
|||
|||
Db 616 TTTGAGGCATTTATGCGCTTCATTTCAGAGTACCAGGTTATGAGTACCATTTATTAAGCGTA 675
```

```
QY 226 TyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAsp 245
Db |||||
Db 676 TAGCTATCCGCTGCAAACTCCATTTATTATTATTAAAGAGATAGCTCGATTTTCGGTTG 735
QY 246 GluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAsp 265
Db |||||
Db 736 GANTGGGGATTA-----AGTCAACTCATGTTAAAC-----GAT 768
QY 266 TyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyr 285
Db |||||
Db 769 AATTATAATCTCAATAATAGGCGCTCGAGATTATGCAAAATCATTTGTACAACTTGGTAT 828
QY 286 ArgThrGlyLeuLysAsnLeuAspGluProAsnMethLysTrpSerIlePheAsnAsp 305
Db |||||
Db 829 CGGACGGGTTTACAAAGATTGCAAGGCACCAATGTAGCAGTTGGGTCAATTATATATCGA 888
QY 306 TyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyrAsp 325
Db |||||
Db 889 TTTAGAGAGAAATGACACTAATCTGTATTAGATGTTTGTGCATTTATTTTCAAGTTATGAT 948
QY 326 IleLysArgTyrArgAspSerIleGlyIleGluValLysGlyIleLysAsnGluLeu 345
Db |||||
Db 949 TATCGTAGTTAC-----CCAATGGAGCTAAGGGA-----GAGCTT 984
QY 346 ThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeuArgValGln 365
Db |||||
Db 985 ACGAGAGAAATTTATACGGATCCAGTAGGAGCCTCTTT-----TGGGTGAATAGAGCA 1038
QY 366 ProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPheLysLeuPheSerPhe 385
Db |||||
Db 1039 CCAAACTTCGCATCAATCAATAATACAGTAGTTAGGCAACACACCCC-----TTTACTTGG 1095
QY 386 LeuGluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIle 405
Db |||||
Db 1096 CTAGTTACTTTTAACAGTTAATACAGTCAAGTCAGATCTGGCGAT-----GGAAT 1146
QY 406 SerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArg 425
Db |||||
Db 1147 TCTAATATTATTGGAAATCACATAGTCAACCGTGAGTGMAACCGGAGGTCCAGTCCT 1206
QY 426 ThrGlySerProThr-----ThrLysThrIleArgProPheGluSerTyrLysVal 442
Db |||||
Db 1207 APTCAGAGTCTCTACCTGTGGAAGTACTGTGTACAAAT-----TATCGC--- 1248
QY 443 SerIleValThrAspArgGln-----SerProValSerProIleGlnProHisPhe 460
Db |||||
Db 1249 -----ACGATAAATTTACTTTTAAATCCA-----TTT 1275
QY 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr 476
Db |||||
Db 1276 TTATTAGGTGATATTATATACCATTAATACAGGTATTGTTCTTATCTGGCTAATTTGTTT 1335
QY 477 LeuLysTyrSerAlaGlyGlySerLeuSerAsn-----TyrGlnAsnThr 491
Db |||||
Db 1336 GGAATCTATTACAGTAGATTATAGCAGACTCGTGTATTGAGCTTCTGTATGAGAACCA 1395
QY 492 ThrPhePheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSer 511
Db |||||
Db 1396 AGAGTTTTCGCCGCTTACAAATCATCAAAATTCGTGAATTA-----CCTGAGGTAAAC 1446
QY 512 -----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThr 528
Db |||||
Db 1447 TCGGATAGCCCAACTGCTGCCGACTATAGTATCATGACTATCATGATATCTCAGGTTTGA 1506
QY 529 TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGly-----Val 544
Db |||||
Db 1507 -----ACTGATGTGGGAGAACGGTTCTAGTT 1533
QY 545 LeuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSer-----AspLysIleIle 563
Db |||||
Db 1534 TATGGGTGACACTCTTCAACTGCTACTCGTGAGAAATAAATTATCGTGTAGACGACAGANA 1593
```

```
QY 564 ThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGluGly 583
Db |||||
Db 1594 GTACAACTTCCAGCTGTTAAAGGGAACAAGTCTCAAC---AATTGCCAAGTAGTTAGAGGA 1650
QY 584 ProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIleThr 603
Db |||||
Db 1651 ACTGGATTTTACAGAGGAGACTGGTTGAAGCCTTAATAATATGTCATCATTTTCTCTAGCT 1710
QY 604 CysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGly 623
Db |||||
Db 1711 CTTGTTTC---AGTTCGACTTACACTTACCGCTCGCATTCGTTATGCTGCCGACGA 1767
QY 624 AlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIleProPro 643
Db |||||
Db 1768 GCGCGATCA-----GGTTTTCTCTGTTATATCGGATCAATATGGAGAATTTTCCA 1818
QY 644 GlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn-----AsnLeuGlnTyr 660
Db |||||
Db 1819 ---ACCACACAGATATCGCTTCTCCACAATGTACTCTGCCCCCAAAATGTACCATAC 1875
QY 661 GlyAspPheGlyTyrPheGlnPheProSerThrValThrLeu----- 674
Db |||||
Db 1876 GAGGCTTTTAAGATTGTAGATTTACCTTCTACTGTACTATTAGAAATACTTCTCTCTGCT 1935
QY 675 -----ProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSer 692
Db |||||
Db 1936 TCAACAACTTTTCGACTTGTATTTCCGTTTTCATTTGTGCATTTAGGAATACTCGCAAT--- 1992
QY 693 IleLeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArg 712
Db |||||
Db 1993 ---ATATTAAATGACCGAATTTGAATTTGTTCCCATGAGGGTTCCTTGTTCGAGTACGAA 2049
QY 713 GluLysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsnHisThr 732
Db |||||
Db 2050 ACCAAACAGCAGCTAGAAAAGCAAGCAAGCGGTGAACCATTTGTTTACAGATGGATCG 2109
QY 733 Lys 733
Db |||||
Db 2110 AAA 2112
Db |||||
RESULT 6
US-11-058-727-13
; Sequence 13, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 2022
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (truncated)
; FEATURE:
; NAME/KEY: CDS
```


Db 1705 CATATGCGTAAATTTCA GACATTAACAAGAAATATAGTAGG 1752
Qy 617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
Db 1753 ATTGCGTATGCTTCC-----GCTAATAATACTGAATTTTATATAAATCTTCTGAAGAA 1806
Qy 637 GlyValIleGlyLeuProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn 656
Db 1807 AACGTTAAATCT---CACGCTCAAAAACATGATGAATAGAGGTGAAGCTTTAAACATATAAT 1863
Qy 657 AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValThr 673
Db 1864 AAATTTAATTATGCGACTTGGCCCTTAAATTTACGACAAACCGAACCCTTTCACTTACT 1923
Qy 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693
Db 1924 CTAGGGCT-----ATATTGAAGCGGAAGACCTTTCTTGGAAATGAA 1965
Qy 694 LeuIleIleAspIleGluPheIleProIle 704
Db 1966 GCTTATATAGCGAATCGAATTTATCCAGTA 1998

RESULT 7

US-11-108-389-13
; Sequence 13, Application US/11108389
; Publication NO. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE REFERENCE: 35718/291049
; CURRENT FILING DATE: US/11/108,389
; PRIOR APPLICATION NUMBER: 2005-04-18
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 2022
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (truncated)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2022)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: 1218-2A
US-11-108-389-13
Alignment Scores:
Pred. No.: 3.52e-62 Length: 2022
Score: 678.50 Matches: 212
Percent Similarity: 44.61% Conservative: 123
Best Local Similarity: 28.23% Mismatches: 277
Query Match: 17.54% Indels: 139
Gaps: 31
US-10-783-417-2 (1-735) x US-11-108-389-13 (1-2022)

Qy 5 AsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyrPheProAsnArg 24
Db 10 AATAATCAAAATCAATATGAATATATAGATCGACACCTTCT-----ACTTCT 57
Qy 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
Db 58 GTATCCAAATGATTTCTAACAGATACCTTTTTCGGAATGAGCCACAAATACGCTACAAAT 117
Qy 44 ThrAsnTyrIleGlyLeuTrpLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
Db 118 ATGGATTATAAAGATTATTTAAAAATGCTCGCGGAATGCTAGTAATAACCTCGGTTC 177
Qy 63 PheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIleValSer 82
Db 178 CCTGAAGTACTTGTTCGCGGACAAGAT-----GCAGCTAAGCGCGCAATGATATAGTA 231
Qy 83 GlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
Db 232 GGTAATTTACTATCAGGTTTAGG-----GTCCCAATTTGTT 267
Qy 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGln 122
Db 268 GGGCCGATAGTGAGTCTTTTATCTCAACTTATTGATATTCTGTGCGCTTCAGGGCAA--- 324
Qy 123 AspIleThrValTrpThrGlnPheIleLysMetGlyIlePheValAspThrProLeu 142
Db 325 ---AAGAGTCAATGGGAGATTTTATGGAAACAAGTAGAAGAACTCATAAATCAAAAAATA 381
Qy 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
Db 382 GCAGAATATGCAAGGAATAAAGCGCTTCGGAATTAGAGGATTAGTAATAATTACCAA 441
Qy 163 SerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
Db 442 TTATATCTAACTCGCTTGAAGAAATGGAAGAA----- 474
Qy 183 ProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGluAsnVal 202
Db 475 AATCCAAATGGTTCA-----AGAGCGTTACGAGATGCGAAATCGATTTGAAATCCTG 528
Qy 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
Db 529 GATAGTTTATTTACGCAATACATGCCATCTTTTCGAGTGACAAATTTTGAAGTACCATT 588
Qy 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGlyAlaGlu 242
Db 589 CTTACAGTATATACACAGCGACCACTTCATTCTGTTTAAAGGAGCGCTTCAATT 648
Qy 243 LeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
Db 649 TTTGGAGAAGAATGGGATGG-----TCTACAACCACT 681
Qy 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
Db 682 ATTAATAACTATTATGATCGTCAAACTTACTCGACAATATTCTGATCACTGTGTA 741
Qy 283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle 302
Db 742 AAGTGTATGAAACTGTTTACCAAAATTAAGGCGACGCGCTAACCAATGGGTGCAC 801
Qy 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db 802 TATAACCAATTCGTAGAGAAATGACACTGACGGTTTTAGATGTTGTGCAATTATCCCA 861
Qy 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyIleGluValLysGlyIleLys 342
Db 862 AATTATACACACGACGCTAC-----CCAATGGAACGAAAGCA----- 900
Qy 343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362
Db 901 ---CAACTAACAAAGGAAGTATATACA-----GATCCACTGGGGCGGTGA 942
Qy 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373

Db 232 GGTAAATTACTACAGCTTAGGG-----GTCCCAATTGTT 267
Qy 103 GlyAlaIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGln 122
Db 268 GGGCCGATAGTAGTCTTTATATACCACTTATTGATATTCTGTGGCTTCAGGGAA--- 324
Qy 123 AspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu 142
Db 325 --AAGAGTCAATGGGAAATTTTATGGAACAAGTAGAAGAACTCAATTAATCAAAAATA 381
Qy 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
Db 382 GCAGATATATCAAGGAATAAAGCGCTTCGGAATTAGAAGGATTAGGTAATAATTACCAA 441
Qy 163 SerTyrAsnThrAlaLeuAspAspTTPArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
Db 442 TTATATCTAACTCGCTTGAGAAATGGGAAGAA----- 474
Qy 183 ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal 202
Db 475 AATCCAAATGGTTCA-----AGAGCCTTACGAGATGTGCGAAATCGATTGAAATCCTG 528
Qy 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
Db 529 GATAGTTATTATACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTACCATT 588
Qy 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlyAlaGlu 242
Db 589 CTCTAGTATATGCAATGGCAGCACTTCACTTACTGTATTAAAGGACGGTCAATT 648
Qy 243 LeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
Db 649 TTTGGGAAGAATGGGATGG-----TCAACAACACTACT 681
Qy 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
Db 682 ATTAATAACTATTATGATCGTCAAAATGAACCTTACTGCAGAAATATTCTGATCACTGTGTA 741
Qy 283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle 302
Db 742 AAGTGTATGAAACTGGTTTTAGCAAAATTAAGGCGACGCGCTAAACAAATGGGTGAC 801
Qy 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db 802 TATAACCAATTCGTTAGAGAAATGACACTGGCGGTTTAGATGTTGTGCAATTATCCCA 861
Qy 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyLeuValLysGlyIleLys 342
Db 862 AATTATGACACGACGCTAC-----CCAAATGGAAACGAAGCA----- 900
Qy 343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362
Db 901 ---CAAATAACAGGAAGTATATACA-----GATCCACTGGGCGCGGTA 942
Qy 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
Db 943 AAGTGCTCTCAATTGGTCTCGTATGATCAAAAGCACTTCTTTCGGAGTGATAGATCA 1002
Qy 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
Db 1003 TCCGTTATTGACACCC---CATGTATTGATTATATACGGGACTCACTACAGTGTATACA 1059
Qy 394 GluAsnThrAsnPheGlyAsn---ArgLeuVal-----GlyIleSer 406
Db 1060 CAATCAAGAAAGCAATTTCTCCGCTCGCTATATAGACATTTGGGCTGCATCAAAATAGC 1119
Qy 407 AsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThr 426
Db 1120 TACCATCGTGTAGTAGGGTAGTAAT-----CTTCAACAATGATAGAACTAATCAA 1173
Qy 427 GlySerProThrThrLysThrIleArgProPheGluSerTyrLysValSerIleValThr 446

Db 1174 AATCTACACAGCACTAGTACCTTT---GATTTTACGAATTATGATATTTACAAGACTCTA 1230
Qy 447 AspArgGlnSerProProValSerProIleGlnPro-----HisPheIle 461
Db 1231 TCAAGAGTACAGTACTCTTGTATTTGTTTACCTGTGTATACGTATATATTTTTTGA 1290
Qy 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerSerAsnAsn-----ThrLeuLys 478
Db 1291 ATGCCAAGATCGAGTCTTTCATGCTAAACCAATTTGAATAATACCAAGAACGCTTAAAG 1350
Qy 479 TyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArg 498
Db 1351 TAT-----AATCCAGCTTTCCAAAGATATTTATAGCGAGTACAAGA 1389
Qy 499 LysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSer 518
Db 1390 GATTCGGAATTAGNAATTCCTCCAGAACTTCAGATCAACCAATTTATAGTCAATAGC 1449
Qy 519 HisIleLeuSerHisPhe---SerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeu 537
Db 1450 CATAGATTATGTATCATATCACAGTATTCCCGCGACGGTAACTACCTACCGGATTA----- 1503
Qy 538 GlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAla 557
Db 1504 -----GTACCTGTATTTCTTGACACATCGAAGTCGAGATTTTAAACAATAACA 1551
Qy 558 IleSerAspLysIleIleThrMetIleProAlaIleLys---GlyAsnAsnLeuAspThr 576
Db 1552 ATATATTCAGATAAAATCACTCAAAATTCGCGCGTTAAATGTTGGGATAAATTTACCGTT 1611
Qy 577 AsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuVal---TyrLeuGln 595
Db 1612 ---GTTCCAGTGGTAAAGACAGGACATACAGAGGGGATTTATTACAGTATAATAGA 1668
Qy 596 SerGlnGly-----ArgLeuGluIleThrCysGluThrPro 607
Db 1669 AGTACTGGTCTGTAGGAACCTTATTCTAGCTCGATATGCGCTAGCATTAGAA----- 1722
Qy 608 AsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThr 627
Db 1723 ---AAGCAGGGAATATCGTGAAGACTGAGATATGCTACTGATGCA----- 1767
Qy 628 LeuProAsnIleSerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsn 647
Db 1768 -----GATATTGTATTGATGTAACGATGCTCAGATTCAGATGCTCAAAACCAATGAAC 1821
Qy 648 AsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGln 667
Db 1822 CCAGGT-----GAGGATCTGCATCTAAACCTTTTAAAGTTGCAGAT 1863
Qy 668 PheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAla 687
Db 1864 GCTATCAACAATTAATTTAGCAACAGATGTCGTAGCAATTTGAACATAATTTAGGT 1923
Qy 688 AspValSerAsnSer-----IleLeuIleIleAspLysIleGluPheIlePro 703
Db 1924 GAAGACCTTAATCAACATTTATCTGTTAGTTTACGTTGACCGCAATCAATTCATCCA 1983
Qy 704 IleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLys 723
Db 1984 GTAGATGAGACATAT-----GAAGCGGAACAAGATTTTAGAAGCGACGAAGAAAGCA 2034
Qy 724 IleAsnThrPhePheThrAsn 730
Db 2035 GTGAATCCCTTGTTTACGAAT 2055

RESULT 9

US-11-108-389-1
; Sequence 1, Application US/11108389
; Publication NO. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan

382	Db	GCAGAAATATGCAAGGAATAAAGCGCTTTTCGGAAATTAGAAGGATTAGAGTAATAATATACCAA	441
163	Qy	SerTyrAsnThrAlaLeuAspAspTirArgLysLeuLysArgLeuGlnAlaProGlyLeu	182
442	Db	TTATATCTAATCGCTTGGAGATGGGAAGA	474
183	Qy	ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal	202
475	Db	AATCCAAATGGTTCA-----AGAGCCTTACGAGATGCGAAATCGATTGGAATCCTG	528
203	Qy	HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu	222
529	Db	GATAGTTTATTTACGCCAATATATGCCACTTTTAGAGTGACAAATTTTGAAGTACCAATTC	588
223	Qy	LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGlu	242
589	Db	CTTACTGTATATGCATGGCAGCCAACTTCATTACTGTATTATTAAGGACGGCTCAATT	648
243	Qy	LeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr	262
649	Db	TTTGGAGAAATGGGGATGG-----TCAACAACACTACT	681
263	Qy	SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla	282
682	Db	ATTAATAACATATTATGATCGTCAATGAAATTTACTGACAAATTTCTGATCTGTGTA	741
283	Qy	AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle	302
742	Db	AGTGGTATCGAACTGGTTTAGCAAAATTAAGGCACGCGCTAAACATGGGTTGAC	801
303	Qy	PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer	322
802	Db	TATAACCAATTCCGTAGAGAAATGCACACTGGCGGTTTGTAGATGTTGTTCATTATCCCA	861
323	Qy	LeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIleLys	342
862	Db	AATTATGACACACGCACGTAC-----CCAATGGAAACGAAGA-----	900
343	Qy	AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu	362
901	Db	---CAACTAACAGGGAAGTATATACA-----GATCCACTGGGGCGGTA	942
363	Qy	ArgValGln-----ProAsnLeuAlaThrMetGluTyr	373
943	Db	AACGTGCTTCAATTTGGTTCTCGTATGACAAAGCACCTTCTTCGGAGTGATAGATCA	1002
374	Qy	AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyrThr	393
1003	Db	TCCGTTATTTCGACCAACC-----CATGTATTTGATTATATATAACGGGACTCACAGTGTATACA	1059
394	Qy	GluAsnThrAsnPheGlyAsn---ArgLeuVal-----GlyIleSer	406
1060	Db	CAATCAGAAGCAATTTCTCCGCTCGCTATATTAAGACATTTGGGCTGTCATCAAAATAGC	1119
407	Qy	AsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyIleArgThr	426
1120	Db	TACCATCGTCTCAGTAGGGTAGTAAAT-----CTTCAACAAATGATGGAACCTAATCAA	1173
427	Qy	GlySerProThrThrLysThrIleArgProPheGluSerTyrLysValSerIleValThr	446
1174	Db	AATCTACACAGCACTAGTACTTTT---GATTTTACGAAATTTATGATATTTTACAAGACTCTA	1230
447	Qy	AspArgGlnSerProProValSerProIleGlnPro-----HisPheIle	461
1231	Db	TCAAGATGCAGTACTCTCTGTATTTGTTTATCCCTGGTTATACGTATATATTTTGGGA	1290
462	Qy	IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys	478
1291	Db	ATGCCAGAAGTCAGTGTGTTTTTCATGGTAAACCAATTTGAATAATACCAGAAAGACGTTAAAG	1350
479	Qy	TyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArg	498
1351	Db	TAT-----AATCCAGTTTCCAAAGATATTTATAGCGAGTACAAAGA	1389

```
QY 499 LysLysAspCysAsnLeuValIleAAspProGlyCysSerProAsnPheAsnAsnTyrSer 518
   ::      |||      ::      |||      ::      |||      ::      |||
Db 1390 GATTTCGAATTAGAAATACCTCCAGAAATTCAGATCAACCAAAATATAGTATATAGC 1449
   ::      |||      ::      |||      ::      |||      ::      |||
QY 519 HisIleLeuSerHisPhe---SerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeu 537
   |||      |||      |||      |||      |||      |||      |||      |||
Db 1450 CATAGATTATGTCATATCACAAGTATTCGCCGACGGGTAACTACCGGATTA----- 1503
   |||      |||      |||      |||      |||      |||      |||      |||
QY 538 GlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAla 557
   |||      |||      |||      |||      |||      |||      |||      |||
Db 1504 -----GTACCTGTATTTCTTGACACATCGAAGTCAGATTTAAACATACA 1551
   |||      |||      |||      |||      |||      |||      |||      |||
QY 558 IleSerAspLysIleIleThrMetIleProAlaIleLys---GlyAsnAsnLeuAspThr 576
   |||      |||      |||      |||      |||      |||      |||      |||
Db 1552 ATATATTTCAGATAAAATCACTCAAAATCCCGCGGTAAATGTTGGGATAATTACCGTTT 1611
   |||      |||      |||      |||      |||      |||      |||      |||
QY 577 AsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuVal---TyrLeuGln 595
   |||      |||      |||      |||      |||      |||      |||      |||
Db 1612 ---GTTCCAGTGGTAAAGGACCAAGGACATACAGAGGGGATTTATTACAGTATAATAGA 1668
   |||      |||      |||      |||      |||      |||      |||      |||
QY 596 SerGlnGly-----ArgLeuGluIleThrCysGluThrPro 607
   |||      |||      |||      |||      |||      |||      |||      |||
Db 1669 AGTACTGGTCTCTAGGAACCTTATTCTAGCTCGATATGCCCTAGCAATTAGAA----- 1722
   |||      |||      |||      |||      |||      |||      |||      |||
QY 608 AsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThr 627
   |||      |||      |||      |||      |||      |||      |||      |||
Db 1723 ---AAAGCAGGGAATATCGTGAAGACTGAGATATGCTACTATGCA----- 1767
   |||      |||      |||      |||      |||      |||      |||      |||
QY 628 LeuProAsnIleSerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsn 647
   |||      |||      |||      |||      |||      |||      |||      |||
Db 1768 -----GATATTGTAATTGATGTAAACAGTCTCAGATTCCAGATGCCAAAACAAATGAAC 1821
   |||      |||      |||      |||      |||      |||      |||      |||
QY 648 AsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGln 667
   |||      |||      |||      |||      |||      |||      |||      |||
Db 1822 CCAGGT-----GAGGATCTCAGATCTAAAACTTTTAAAGTTGCAGAT 1863
   |||      |||      |||      |||      |||      |||      |||      |||
QY 668 PheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAla 687
   |||      |||      |||      |||      |||      |||      |||      |||
Db 1864 GCTATCACACATTAATTTAGCAACAGATAGTTCGTAGCATTTGAACATAATTAGGT 1923
   |||      |||      |||      |||      |||      |||      |||      |||
QY 688 AspValSerAsnSer-----IleLeuIleIleAspLysIleGluPheIlePro 703
   |||      |||      |||      |||      |||      |||      |||      |||
Db 1924 GAAGACCCTAAATCAACATTTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCA 1983
   |||      |||      |||      |||      |||      |||      |||      |||
QY 704 IleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLys 723
   |||      |||      |||      |||      |||      |||      |||      |||
Db 1984 GTAGATGAGACATAT-----GAAGCGGAACAAGATTTAGAACGACGCAAGAAAGCA 2034
   |||      |||      |||      |||      |||      |||      |||      |||
QY 724 IleAsnThrPhePheThrAsn 730
   |||      |||      |||      |||      |||      |||      |||      |||
Db 2035 GTGAATGCCCTTGTTTAGCAAT 2055
   |||      |||      |||      |||      |||      |||      |||      |||
RESULT 10
US-11-058-727-17
; Sequence 17, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
```

```
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 4874
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Genomic DNA 1218-1
US-11-058-727-17

Alignment Scores:
Pred. No.: 3,51e-61 Length: 4874
Score: 675.00 Matches: 216
Percent Similarity: 44.46% Conservative: 125
Best Local Similarity: 28.16% Mismatches: 300
Query Match: 17.45% Indels: 126
DB: 32

US-10-783-417-2 (1-735) x US-11-058-727-17 (1-4874)
QY 5 AsnAspAsnAsnGluTyrGluIleAspSerHisThrSerProTyrPheProAsnArg 24
   |||      |||      |||      |||      |||      |||      |||      |||
Db 740 AATTAACAAATGAATATGAATATATAGTAGCGACACCTTCT-----ACTTCT 787
   |||      |||      |||      |||      |||      |||      |||      |||
QY 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
   |||      |||      |||      |||      |||      |||      |||      |||
Db 788 GTATCCATGATCTTAACACATACCTTTTGGGATGAGCCACAAATCGCTACAAAT 847
   |||      |||      |||      |||      |||      |||      |||      |||
QY 44 ThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
   |||      |||      |||      |||      |||      |||      |||      |||
Db 848 ATGGATTATAAAGATATTATTAATAATGCTCGCGGAAATGCTAGTAATACCTGGTTCA 907
   |||      |||      |||      |||      |||      |||      |||      |||
QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSer 82
   |||      |||      |||      |||      |||      |||      |||      |||
Db 908 CTGGAAGTACTGTGTAGCGGACCAAGAT-----GCAGCTAAGCGCCCAATTGATATAGA 961
   |||      |||      |||      |||      |||      |||      |||      |||
QY 83 GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
   |||      |||      |||      |||      |||      |||      |||      |||
Db 962 GGTAAATTACTATCAGGTTTAGGG-----GTCCCATTTGTT 997
   |||      |||      |||      |||      |||      |||      |||      |||
QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGlyGluGln 122
   |||      |||      |||      |||      |||      |||      |||      |||
Db 998 GGGCCGATAGTAGTCTTTATCTCAACTTATTGATATTCTGTGCGCTTCAGGGGAA--- 1054
   |||      |||      |||      |||      |||      |||      |||      |||
QY 123 AspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu 142
   |||      |||      |||      |||      |||      |||      |||      |||
Db 1055 ---AAGACTCAATGGGAAATTTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATA 1111
   |||      |||      |||      |||      |||      |||      |||      |||
QY 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
   |||      |||      |||      |||      |||      |||      |||      |||
Db 1112 GCAGATATCAAGGAATAAGCGCTTTCGGAATTAGAGGATTAGGTAATAATTACCAC 1171
   |||      |||      |||      |||      |||      |||      |||      |||
QY 163 SerTyrAsnThrAlaLeuAspAspTyrTrpArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
   |||      |||      |||      |||      |||      |||      |||      |||
Db 1172 TTATATCTAACTGCGCTTGAAGAATGGAAGAA----- 1204
   |||      |||      |||      |||      |||      |||      |||      |||
QY 183 ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal 202
   |||      |||      |||      |||      |||      |||      |||      |||
Db 1205 AATCCAAATGGTTCA-----AGAGCCTTACGAGATGCGAAATCGAATTTGAAATCCTG 1258
   |||      |||      |||      |||      |||      |||      |||      |||
QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
   |||      |||      |||      |||      |||      |||      |||      |||
Db 1259 GATAGTTTATTACGCAATATATGCCATCTTTTAGAGTGACAAATTTGAAGTACCATTC 1318
   |||      |||      |||      |||      |||      |||      |||      |||
QY 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGlyAlaGlu 242
   |||      |||      |||      |||      |||      |||      |||      |||
```

1319 CTTACTGTATATGCAATGGCGAGCCAACTTTCATTTACTGTTATTAAAGGACGCGTCAATT 137

243 LeuAlaAspGluThrProAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262

1379 TTTGGAGAAGAATCGGGATGG-----TCACAACACTACT 1411

263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282

1412 ATTAATAACTATTATGATCGTCAAACTGAACTTACTGCGAAGTATCTGATCACTGTGTA 1471

283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTyrSerIle 302

1472 AAGTGGTATGAACCTGGTTTACGCAAAATTAAGACGCGCTAAACATGGGTGGAC 1531

303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322

1532 TATAACCAATTCGCTAGAGAAATGACACTGGCGGTTTATGATGTTGTCATTTATCCCA 1591

323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyLeuValLysGlyIleLys 342

1592 AATTATGACACACGCGCTAC-----CCAATGGAAACGAAAGCA----- 1630

343 AsnGlnLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362

1631 ---CAACTAACAGGAAGTATATACA-----GATCCACTGGGCGCGGT 1672

363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373

1673 AACGTGCTTCAATTTGCTTCTGGTATGACAAAGCACCTTCTTCGGAGTGATAGATCA 1732

374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393

1733 TCGGTATTTCGACCACCC---CATGTATTTCATTATATAACGGGACTCACAGTGTATACA 1789

394 GluAsnThrAsnPheGlyAsn---ArgLeuVal-----GlyIleSer 406

1790 CAATCAAGAACGATTTCTCCGCTCGCTATATAGACATTTGGGCTGTGCATCAAAATAGC 1849

407 AsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyLeuArgThr 426

1850 TACCATCGTGCAGTAGGGTAGTAAT-----CTTCAACAATGATAGACTAATCAATCA 1903

427 GlySerProThrThrLysThrIleArgProPheGluSerTyrLysValSerIleValThr 446

1904 AATCTACAGCACTAGTACTCTT---GATTTTACGAATTATGATATTTACAAGACTCTA 1960

447 AspArgGlnSerProProValSerProIleGlnPro-----HisPheIle 461

1961 TCAAAGGATCGACTCTTGTATTTGTTTACCCTGTTATACGTATATATTTTTTGG 2020

462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478

2021 ATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTGAATATATACCAGAAAGAGCTTAAG 2080

479 TyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArg 498

2081 TAT-----AATCCAGTTTCCAAAGATATATAGCGAGTACAAGA 2119

499 LysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSer 518

2120 GATTCGGAATTAGAATTTACCTCCAGAAACTTCAGATCAACCAATTTATGATGATATAGC 2179

519 HisIleLeuSerHisPhe---SerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeu 537

2180 CATAGATTATGTCATATACAAAGTATTTCCCGGACGGGTAACTACCGGATTA----- 2233

538 GlnIleLeuAspThrGlyValLeuGlyTyrHisSerSerValAspArgTyrAsnAla 557

2234 -----GTAACCTGTATTTCTTGGACACATCGAAGTCGAGTTTAAACNATACA 2281

558 IleSerAspLysIleIleThrMetIleProAlaIleLys---GlyAsnAsnLeuAspThr 576

2282 ATATATTACAGATAAAATCACTCAAAATCCGCGCGTTAAATGTTGGGATATATTTACCGTTT 2341

```

Qy 577 AsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuVal-----TyrLeuGln 595
Db 2342 ---GTTCCAGTGGTAAAGGACCAGGACATACACAGGAGGGATTTATTATACAGTATAATAGA 2398
Qy 596 SerGlnGly-----ArgLeuGluIleThrCysGluThrPro 607
Db 2399 AGTACTGGTTCTGTAGGAACCTTATTCTAGCTCGATATGCCCTAGCATAGAA----- 2452
Qy 608 AsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThr 627
Db 2453 ---AAAGCAGCGGAAATATCTGTGAAGACTCAGATATGCTACTGATGCA----- 2497
Qy 628 LeuProAsnIleSerLeuThrIleProGlyValIleGlyLeProGlnArgLeuAsn 647
Db 2498 -----GATATTGTATTGCAATGATAAAGCATGCTCAGATTCAGATGCCAAACAAATGAAC 2551
Qy 648 AsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGln 667
Db 2552 CCAGGT-----GAGGATCTGCATCTAAAACTTTAAAGTTGCAGAT 2593
Qy 668 PheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAla 687
Db 2594 GCTATCACACAACTAAATTTAGCAACAGATAGTTCGTAGCATTTGAAACATAATTTAGGT 2653
Qy 688 AspValSerAsnSer-----IleLeuIleIleAspLysIleGluPheIlePro 703
Db 2654 GAAGACCTCAATTTCAACATTTATCTGTTATAGTTTACGTTGACCGAATTCGAAATTCATCCCA 2713
Qy 704 IleThrSerSerMethHisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLys 723
Db 2714 GTAGATGAGACATAT-----GAAGCGGAACAGATTTAGACGACGGAAGAAGCA 2764
Qy 724 IleAsnThrPhePheThrAsn 730
Db 2765 GTGAATGCCTTGTGTACGAAT 2785

RESULT 11
US-11-108-389-17
; Sequence 17, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 4874
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Genomic DNA 1218-1
US-11-108-389-17

```

Alignment Scores:				
Pred. No.:	3.51e-61	Length:	4874	
Score:	675.00	Matches:	216	
Percent Similarity:	44.46%	Conservative:	125	
Best Local Similarity:	28.16%	Mismatches:	300	
Query Match:	17.45%	Indels:	126	
DB:	7	Gaps:	32	
US-10-783-417-2 (1-735) x US-11-108-389-17 (1-4874)				
QY	5	AsnAspAsnAsnGluTyrGluIleAspSerHisThrSerProTyrPheProAsnArg	24	
DB	740	AAATAATCAAATGAATATGAATATATAGATCGGACACTTCTT-----ACTTCT	787	
QY	25	AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn	43	
DB	788	GTATCCAAATGATCTAACAGATACCTTTTCCGAATGAGCCACAAATCGCTACAAAT	847	
QY	44	ThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn	62	
DB	848	ATGGATATTAAGATATATTTAAAAATGTCTCGGGAAATCTAGTGAATACCTCGTGTCA	907	
QY	63	PheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIleValSer	82	
DB	908	CCTGAAGTACTGTTACGGGACAAGAT-----GCAGTAAGCGCCGAATGATATAGTA	961	
QY	83	GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIle	102	
DB	962	GGTAATAATTAATCATCAGGTTTAGGG-----GTCCCAATTTGTT	997	
QY	103	GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGlyGluGln	122	
DB	998	GGCCGATATGAGTCTTTATATCTCAACTATTATGATATCTGTGGCCTTCAGGGGAA	1054	
QY	123	AspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu	142	
DB	1055	---AAGAGTCAATGGGAAATTTTATCGCAACAGTAGAGAACTCATTAATCAAAAAATA	1111	
QY	143	ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln	162	
DB	1112	GCAGAATATGCAAGGAATAAAGCGCTTTCGGAATTAGAAGATTAGGTAATAATTACCAA	1171	
QY	163	SerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaProGlyLeu	182	
DB	1172	TTATATCTAACTGCGCTTGAAGATGCGGAAGAA-----	1204	
QY	183	ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal	202	
DB	1205	AAATCCAAATGGTTCA-----AGAGCCTTACGAGATGTGGAAATCGATTTGAATCCTG	1258	
QY	203	HisAsnAspPheLeuArgGluIleProGlyPheGlnLeuGluTyrLysThrLeuLeu	222	
DB	1259	GATAGTTTATTACGCAATATATGCCATCTTTTAGAGTGACAAATTTGAAGTACCAATTC	1318	
QY	223	LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlyAlaGlu	242	
DB	1319	CTTACTGTATATGCAATGGCAGCCCACTTCATTACTGTTATTAAAGGACGGGTCAATT	1378	
QY	243	LeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr	262	
DB	1379	TTTGGAGAAGATGGGATGG-----TCACAACTACT	1411	
QY	263	SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla	282	
DB	1412	ATTAATAACATTAATATGATCGTCAAAATGAAATCTTACTGCAGAAATATCTGATCTGTGA	1471	
QY	283	AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle	302	
DB	1472	AAGTGTATGAACCTGTTTAGCMAAATTAAGGCGACGAGCGCTAAACCAATGGGTGCAC	1531	
QY	303	PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer	322	

```

QY 648 AsnThrPheSerGlyThrAsnThrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGln 667
Db 2552 CCAGGT-----GAGATCTGACATCACTAAACCTTTTAAAGTTGCGAGAT 2593
QY 668 PheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAla 687
Db 2594 GCTATCACAACTTAATTTAGCAACAGATATTCGCTAGCATGAAACATAATTTAGGT 2653
QY 688 AspValSerAsnSer-----IleLeuIleIleAspLysIleGluPheIlePro 703
Db 2654 GAAGACCTAATTCACATATCTGTTATAGTTTACGTTGACCGAATCGAATTCATCCCA 2713
QY 704 IleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLys 723
Db 2714 GTAGATGAGACATAT-----GAAGCGGAACAGATTTAGAAAGCGGCAAGAAAGCA 2764
QY 724 IleAsnThrPhePheThrAsn 730
Db 2765 GTGAATGCCTTGTGTACGAAT 2785

RESULT 12
US-11-091-643-3
; Sequence 3, Application US/11091643
; Publication No. US20050246789A1
; GENERAL INFORMATION:
; APPLICANT: TANAKA, Masao
; APPLICANT: YOKOYAMA, Tomoko
; APPLICANT: AOYAGI, Moriochi
; APPLICANT: HASEGAWA, Makoto
; APPLICANT: EHARA, Gaku
; APPLICANT: KIMURA, Masaharu
; APPLICANT: NISHIHASHI, Hideji
; TITLE OF INVENTION: Polypeptide having larvae growth inhibiting or
; TITLE OF INVENTION: insecticidal effect on scarabaeidae insects and
; TITLE OF INVENTION: poynuclotide encoding the same
; FILE REFERENCE: OP1335
; CURRENT APPLICATION NUMBER: US/11/091,643
; CURRENT FILING DATE: 2005-03-29
; PRIOR APPLICATION NUMBER: JP 2001-115754
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: JP 2001-203463
; PRIOR FILING DATE: 2001-07-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 4359
; TYPE: DNA
; ORGANISM: Bacillus popilliae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (282)..(4229)
US-11-091-643-3

Alignment Scores:
Pred. No.: 1,3e-60 Length: 4359
Score: 669,00 Matches: 230
Percent Similarity: 43.47% Conservative: 106
Best Local Similarity: 29.75% Mismatches: 301
Query Match: 17.29% Indels: 136
DB: 9 Gaps: 34

US-10-783-417-2 (1-735) x US-11-091-643-3 (1-4359)

QY 1 MetAsnGln-----AsnAsnAspAsn-----AsnGluTyr 10
Db 336 ATGAATCAATATCATACCAAAACGATACAAAGTTTACAAACCAAGTGGAAATGAAATG 395
QY 11 GluIleIle-----AspSerHisThrSerProTyrPheProAsnArgAsnSerAsnAsp 28
Db 396 CAATATCACTCACTTCAAGTAAGCGCTTTACTTTACAGTCCCAAT----- 440
QY 29 SerArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsnThrAsnTyrLysGlu 48

```

```

Db 441 ---AAGTATCGGTATGTACCGATCCCAATGTCTACGAGCGGTGGAGTTATATAAAT 497
QY 49 TrpLeuAsnMetCysGlnGlyAsnThrGlnTyrGlyAspAsnPheGluThrPheAlaSer 68
Db 498 TGGTTGGATATGTGTACAGGACA-----GGCGACACACAGAAAGCCCGAAACTGTCT 548
QY 69 AlaAspThrIleAlaAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAlaGly 88
Db 549 GCTATTTTCAAAAGGTGCTTTTCTGCTGCAATTAATAAGC----- 590
QY 89 IleGlyGlyLeuThrSerIleSerGlyProIleGlyIle-----IleGly 103
Db 591 -----ACGGGCTTCTGGCTTACTAGGTGTCGTTGGCATCAAAATCGGG 638
QY 104 AlaIleIleSerPheGlyThrLeuLeuThrValPheThrProAlaGlyGluGlnAsp 123
Db 639 GCATTT-----TATACCTTCTTATTAATACCTTATGGCCTGCA-----AGC 680
QY 124 LysThrValThrThrGlnPheIleIleMetGlyGluIlePheValAspThrProLeuThr 143
Db 681 AATACTCAATGGGAGCAGTTTTATAGCACATGTGGAAGAACTCATAAATGCCAAACTAACA 740
QY 144 GluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGlnSer 163
Db 741 GATCATGTGAAGAAATTCGGCACTTACAAAATTAATGTTTACGPAATAACATAGAGATA 800
QY 164 TyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaProGlyLeuPro 183
Db 801 TATAACGAAGCTTTAATAGTTTGGAAACAA-----GATCCT 836
QY 184 ProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHis 203
Db 837 AACAAATAGCAAACTAAAGAT-----CATGTAGCAAGTAAATTCGTAGTCTAAAT 887
QY 204 AsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 223
Db 888 AGCCAATTCGAAGAAATATATTTCCCAATTTAAAGAAAGAGTTTGAAGTTCAAATTTATTA 947
QY 224 ProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeu 243
Db 948 ACTATATATGCACAACTCTGCAATCTTCATCTATTATTATTAAAGAGATTCCTCTTTGAT 1007
QY 244 AlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSer 263
Db 1008 GGTGCATCTTGGGATTT-----GCTCAAGCTACTATT 1040
QY 264 AspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAlaAsn 283
Db 1041 GACAAATAATTACATTCGCCAAATAAGGAAACCCGACAGATATGCAAAATCATTTGTACCACT 1100
QY 284 ThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysThrSerIlePhe 303
Db 1101 TGGTATCAGACGGGTTTACAAAGATTGCAAGGCACTACTGCTAGCAGTTGGCTCTCTTAT 1160
QY 304 AsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeu 323
Db 1161 CATAGATTAGACAGAGAAATGACACTAACAGTATTGGATATTTGGCATTAATTTTCAAAT 1220
QY 324 TyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIleLysAsn 343
Db 1221 TATGATCCCGTAGTTAC-----CCACTGGAGGTAAGGGGA----- 1256
QY 344 GluLeuThrArgGluIleTyrThrThrGluIleAsn-----PheAspArg 358
Db 1257 GAGTCTACGAGAGAAATTTATACGATCCCGGATCCCGGCTACTAACTGGATAGATCGA 1316
QY 359 LeuProGlnLeuArgValGlnProAsnLeuAla---ThrMetGluTyrAsnLeuThrArg 377
Db 1317 GCACCATCTTCGAGAAATAGAAATCTAGTAATTAGGGCACCAGAACTGTTACTTGG 1376
QY 378 AlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheThrGluAsnThrAsn 397

```

```
Db 1377 ATATCCGGTGATTTAATAGTATATACAGGTAGATTGTACGGCTATACTGTTAATAACGAT 1436
Qy 398 Phe-----GlyAsnArgLeuValGlyIleSerAsnArg-----AspAlaPro 411
   :::::|||||:::
Db 1437 TATTGGGCACACACAGGCTAGATTTTCTTGAAACCAATGTTATCGTTTGGGGTCTT 1496
Qy 412 ThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySerProThrThr 431
   |||||:::|||||:::
Db 1497 ACCTATGGATCGACGATTATATAAGT-----CGTACAGATTCTATTCCCATG 1544
Qy 432 LysThrIleArgProPheGluSerTyrLysValSerIleValThrAspArgGlnSerPro 451
   ::|||:::|||||:::
Db 1545 AATCTATT-----GAGTTTATTCACACTACTGTAGTACTGTGGCTCTGCTGG 1595
Qy 452 ProValSerProIleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGly 471
   |||||:::|||||:::
Db 1596 CCAACTGGC-----GGTTTGTGTGGAGTCGCTTCGGCTAGATTTTTCG 1643
Qy 472 SerSerAsnAsnThrLeuLysTyrSerAlaGlyGlySerLeuSerAsn-----TyrGln 489
   |||||:::|||||:::
Db 1644 AAAAGTCCTAGCACC-----GGTTTATTAGGTGAGCGGGTGTATCAG 1685
Qy 490 AsnThrThrPhePheGlnPheProArgLysAspCysAsnLeuValIleAsp---Pro 508
   |||||:::|||||:::
Db 1686 AATCCAGTATATTT-----TCGAGTTCCACTTTAACTTTAACTTACCT 1730
Qy 509 GlyCys-----SerProAsnPheAsnAsnTyrSerHisIleLeuSerHisPheSer 525
   |||||:::|||||:::
Db 1731 GAGTAGACCAAGATAGCCACTGTCGCGACTATAGTATCATTAATCTGTATCACA 1790
Qy 526 LeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeu 545
   |||||:::|||||:::
Db 1791 GCATTTTCGA-----ACTGGATTGAATGGAACCTGTT-----CCGGTTTTT 1829
Qy 546 GlyTrpThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMet 565
   |||||:::|||||:::
Db 1830 GGACGGTATTCTGCACATGTTAGTCTGTCGACAATAGAAATTCGACGACAAATAACGCA 1889
Qy 566 IleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGluGlyProGly 585
   |||||:::|||||:::
Db 1890 ATCCCGGGCTTTAAGTCAAACTCCCTCGAC---AATTGTCAGTAGTTAGAGGACTGGA 1946
Qy 586 HisThrGlyGlyAsnLeuValTyrLeuGlnIleSerGlnGlyArgLeuGluIleThrCysGlu 605
   |||||:::|||||:::
Db 1947 TTTTACAGGAGGACGCTGTTGAAGACAGTTACTTGTAGTCTTTTGTCTTAACCTACT 2006
Qy 606 ThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThr-----621
   :::::|||||:::
Db 2007 TCATCGAGAGCGGGCCAAATCTTACCGCATCCGGTCTCGTTATGCTGCTGAGTAGATTTA 2066
Qy 622 ---AsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGly 640
   |||||:::|||||:::
Db 2067 ATTATGAGTATATTTCTAATGACCCCTTTTATTTCCAAAGGAATAGTCTTACCAATCA 2126
Qy 641 IleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyr 660
   |||||:::|||||:::
Db 2127 ATGCCACCA-----CTGACCCGAAC-----GTACCTTACGAGCTTTTAATTT 2171
Qy 661 GlyAspPheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIle 680
   |||||:::|||||:::
Db 2172 GCAGATTTTGGTGC-----ACTTTTACAACACAGCTACTGCTAACAAGATAT 2219
Qy 681 ProPheIlePheAsnArgAlaAspValSerAsnSerIleLeuIleAspLysIleGlu 700
   |||||:::|||||:::
Db 2220 ACTTTTCAATTCAT-----ACGGGTGGAGCAGCTATAATTCAGCAATTTGA 2267
Qy 701 PheIleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIle 720
   |||||:::|||||:::
Db 2268 TTGTGTTCCAAATTCAGGGTAGTTTGTTCGAGTAGCAACCAACACACAGCTAGAAAAGCA 2327
Qy 721 GlnThrLysIleAsnThrPhePheThrAsnHisThrLys 733
   :::::|||||:::
Db 2328 AGGAAAGCGGTGAACCATTTGTTTACAGATGGATCGAAA 2366
```

RESULT 13

```
US-11-058-727-73
; Sequence 73, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnall
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; PRIOR FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 2025
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (mutated)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2025)
US-11-058-727-73

Alignment Scores:
Pred. No.: 6,45e-60 Length: 2025
Score: 657.50 Matches: 211
Percent Similarity: 44.67% Conservative: 120
Best Local Similarity: 28.48% Mismatches: 293
Query Match: 16.99% Indels: 117
DB: 31 Gaps: 31

US-10-783-417-2 (1-735) x US-11-058-727-73 (1-2025)
Qy 5 AsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyrPheProAsnArg 24
   |||||:::|||||:::
Db 10 AATAATCAAAATGAATATGAAATATATAGATCGCACCTTCT-----ACTTCT 57
Qy 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
   |||||:::|||||:::
Db 58 GTATCCAAATGATTCTAACAGATACCTTTTTCGGAATGAGCCACAAATCGCCTACAAAT 117
Qy 44 ThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
   |||||:::|||||:::
Db 118 ATGATATTAAAGATTATTTTAAAAATGTCGCGGAAATGCTAGTGAATACCTCGGTTC 177
Qy 63 PheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSer 82
   |||||:::|||||:::
Db 178 CCTGAAGTACTCTTAGCGGACAAGAT-----GCAGCTAAGCGCGCAATGATATAGTA 231
Qy 83 GlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
   |||||:::|||||:::
Db 232 GGTAAATTTACTACTCAGGTTTAGGG-----GTCCCATTTGTT 267
Qy 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGlyGluGln 122
   |||||:::|||||:::
Db 268 GGGCCGATAGTAGTCTTTTATCTCAACTATTGATATTCTGTGGCCTTCAGGGGAA--- 324
Qy 123 AspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu 142
   |||||:::|||||:::
```


Db 325 ---AAGAGTCAATGGGAAATTTTATGGAACAAGTAGAAGAACTCATTAAATCAAAAAATA 381
Qy 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
Db 382 GCAGATATGCAAGGAATAAAGCCGCTTCGGAATTTAGAAAGATTAGTAAATAATTACCAA 441
Qy 163 SerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
Db 442 TTATATCTAATCGCTTGAAGATGGGAAGAA---AATCCATTCGAAGTCGAGGTTTT 498
Qy 183 ProProSerSerAlaLeuGlnAlaLeuThrLeuLysIleArgPheGluAsnVal 202
Db 499 CGAAGTCGAGTGCCTTACGAGAT-----GTGCGAAATCGATTTGAAATCCTG 546
Qy 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
Db 547 GATAGTTTATTCGCAATATATCCATCTTTTAGAGTGACAAATTTTGAAGTACCAATC 606
Qy 223 LeuProIleTyrAlaGlnAlaAsnPheHisLeuAsnLeuGlnGlnGlyAlaGlu 242
Db 607 CTACTGTATGCAATGGGACGCAACCTTCATTCTGTTATTAAAGGACGCGTCAATT 666
Qy 243 LeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
Db 667 TTTCGAGAAGAATGGGATGG-----TCAACAACACTACT 699
Qy 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
Db 700 ATTAATAACTATTATGATCGTCAAAATGAACCTTACTGCAGAAATATCTGATCACTGTGA 759
Qy 283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysThrIle 302
Db 760 AAGTGGTATGAACCTGTTTAGCAAAATTTAAAGCGACGCGCTAAACAATGGGTGTGAC 819
Qy 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db 820 TATTAACCAATTCGCTAGAGAAATGACACTGCGGTTTTAGATGTTGTTCATTAATCCCA 879
Qy 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyIleGluValLysGlyIleLys 342
Db 880 AATTATGACACACGACGCTAC-----CCAATGGAAACGAAAGCA----- 918
Qy 343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362
Db 919 ---CAACTAACAGGGAAGTATATACA-----GATCCACTGGCGGCGGTGA 960
Qy 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
Db 961 AACGTGCTTCAATGGTTCCTGTATGACAAAGCACCTTCTTTCGGAGTGATAGAATCA 1020
Qy 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
Db 1021 TCCGTTATTCGACCACCC---CATGTATTGATTATATAACGGGACTCACAGTGTATACA 1077
Qy 394 GluAsnThrAsnPheGlyAsn---ArgLeuVal-----GlyIleSer 406
Db 1078 CAATCAAGAAGCATTTCTCCGCTCGTATATAAGACATTGGGCTGGTGCATCAATAAAGC 1137
Qy 407 AsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThr 426
Db 1138 TACCATCGTGTAGTAGGGGTAGTAAT-----CTTCAACAAATGTATGGAACCTAATCAA 1191
Qy 427 GlySerProThrThrLysThrIleArgProPheGluSerTyrLysValSerIleValThr 446
Db 1192 AATCTACACGACTAGTACCTTT---GATTTTACGAATATGATATTATTACAGACTCTA 1248
Qy 447 AspArgGlnSerProProValSerProIleGlnPro-----HisPheIle 461
Db 1249 TCAAGGATGCACTACTCTTGTATATGTTTACCTGTTATACGTATATATTTTTCGA 1308
Qy 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
Db 1309 ATGCCAGAAGTCGAGTTTTTCATGCTAAACCAATTTGAATATACCAAGAAAGACGTTAAAG 1368

Qy 479 TyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArg 498
Db 1369 TAT-----AATCCAGTTTCCAAAGATATATTATACCGAGTACAGA 1407
Qy 499 LysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSer 518
Db 1408 GATTCGGAAATTAGATTACCTCCAGAACTTCAGATCAACCAANATTATCAGTCATATAGC 1467
Qy 519 HisIleLeuSerHisPhe---SerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeu 537
Db 1468 CATAGATTATGTCATATACCAAGTATTCCCGCGACGGGTAAACACTACCGGATTA----- 1521
Qy 538 GlnIleLeuAspThrGlyValLeuGlyTyrPheHisSerSerValAspArgTyrAsnAla 557
Db 1522 -----GTACCTGTATTCTTGGACACATCGAAGTGCAGATTATAACAATACA 1569
Qy 558 IleSerAspLysIleIleThrMetIleProAlaIleLys---GlyAsnAsnLeuAspThr 576
Db 1570 ATATATTCAGATAAATCACTCAAAATTCGGCCGTTAAATGTGGGATTAATTTACCGTTT 1629
Qy 577 AsnSerLysValIleGluGlyProGlyHisThrGlyAsnLeuVal---TyrLeuGln 595
Db 1630 ---GTTCCAGTGTAAAGGACGACATACAGAGGGGATTATTACAGTATAATAGA 1686
Qy 596 SerGlnGly-----ArgLeuGluIleThrCysGluThrPro 607
Db 1687 AGTACTGGTCTGTAGGAACCTTATTCTAGCTCGATATGGCTAGCATTAGAA----- 1740
Qy 608 AsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThr 627
Db 1741 ---AAAGCAGGAAATATCGTGTAAAGACTGAGATATGCTACTGATGCA----- 1785
Qy 628 LeuProAsnIleSerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsn 647
Db 1786 -----GATATTGTATTGCATGTAAACGATGCTCAGATTCAGATGCCAAAAACAATGAAC 1839
Qy 648 AsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGln 667
Db 1840 CCAGGT-----GAGGATCTGACATCTAAAAACTTTTAAAGTTGCAGAT 1881
Qy 668 PheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAla 687
Db 1882 GCTATCAACAGATTAATTTAGCAACAGATAGTTCGGTTGCGAGTTAAACATATTTAGGT 1941
Qy 688 AspValSerAsnSer-----IleLeuIleIleAspLysIleGluPheIlePro 703
Db 1942 GAAGACCTTAATTCACATTAATTCGTATAGTTTACGTTGACCGAATCGAATTCATCCA 2001
Qy 704 Ile 704
Db 2002 GTA 2004

RESULT 14

US-11-058-727-79
; Sequence 79, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herzmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnall
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786

```

; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 79
; LENGTH: 2025
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (mutated)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2025)
; US-11-058-727-79

```

```

Alignment Scores:
Pred. No.: 6,45e-60 Length: 2025
Score: 657.50 Matches: 212
Percent Similarity: 44.80% Conserved: 120
Best Local Similarity: 28.61% Mismatches: 292
Query Match: 16.99% Indels: 117
DB: 31 Gaps: 31

```

US-10-783-417-2 (1-735) x US-11-058-727-79 (1-2025)

```

Qy 5 AsnAspAsnGluThrGluLeuIleAspSerHisThrSerProTyrPheProAsnArg 24
Db 10 AATAATCAAAATGAATATGAATATATAGATCGCACACCTTCT-----ACTTCT 57
Qy 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
Db 58 GTATCCAAATGATCTTAACAGATACCCCTTTTGGCAATGAGCCAAACAAATGGCTTACAAAAT 117
Qy 44 ThrAsnTyrLysGluThrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
Db 118 ATGGATTATAAGATTATTTAAAAATGTCTGGGGAATGCTAGTGAATACCCCTGGTTCA 177
Qy 63 PheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSer 82
Db 178 CTGAAGTACTTGTAGCGGACCAAGAT-----GCAGCTAAGCGCGCAATTGATATAGTA 231
Qy 83 GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
Db 232 CGTAAATTAATCTACAGTTAGG-----GTCCCAATTTGTT 267
Qy 103 GlyAlaIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGlyGluGln 122
Db 268 GGGCGGATAGTAGTCTTTTACTCACTTATGATATCTGTGGCTTCAGGGGAA--- 324
Qy 123 AspLysThrValThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu 142
Db 325 ---AAGAGTCATGGGAATTTTATGGACAAGTAGAAGCACTCATTAATCAAAAATA 381
Qy 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
Db 382 GCAGATATCAAGGAATAAAGCGCTTCGGAATTAGAGGATTAGTAATAATTACCAA 441
Qy 163 SerTyrAsnThrAlaLeuAspAspThrArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
Db 442 TTATATCTAACTCGCGTTGAAGAATGGGAAGAA---AATCCATTTCCAGTCAGGTCCA 498
Qy 183 ProProSerSerAlaLeuGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal 202
Db 499 ATGGTTCGGGCGCTTACAGAT-----GTGCGAATCGAATTTGAATCCCTG 546
Qy 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
Db 547 GATAGTTATTTACGCAATATATGCCATCTTTAGAGTAGCAAAATTTGAAGTACCATTC 606
Qy 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGlyAlaGlu 242
Db 607 CTTACTGTATATGCAATGGCAGCAACCTTTCATTTACTGTATTATAAAGGAGCGGTCAATT 666

```

```

Qy 243 LeuAlaAspGluThrPheAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
Db 667 TTTGGAGAGAAATGGGGATGG-----TCAACAACACTACT 699
Qy 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
Db 700 ATTAATAACTATTATGATCGTCAAAATGAACTTACTCGAGAAATTTCTGACACTGTGTA 759
Qy 283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle 302
Db 760 AAGTGGTATGAACCTGTTTAGCAAAATTAAGAGCGCAGCGCTAAACAATGGTTGAC 819
Qy 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db 820 TATAACCAATTCGGTAGAGAAATGACACTGGCGGTTTGTAGATGTTGTGCTATTATCCCA 879
Qy 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIleLys 342
Db 880 AATTATGACACACGACGCTAC-----CCAATGGAAACGAAAGCA----- 918
Qy 343 AsnGluLeuThrArgGluIleTyrThrGluIleAsnPheAspArgLeuProGlnLeu 362
Db 919 ---CAACTAACAGGGAAGTATATACA-----GATCCACTGGGCGCGTA 960
Qy 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
Db 961 AACGTGCTTCAATTTGGTCTCGTATGACAAAGCACCTTCTTTCGGAGTGATAGATCA 1020
Qy 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
Db 1021 TCGGTTATTTCGACACCC-----CATGTATTGATTATATATACGGGACTCACAGTGTATACA 1077
Qy 394 GluAsnThrAsnPheGlyAsn---ArgLeuVal-----GlyIleSer 406
Db 1078 CAATCAAGAGCAATTTCTTCGGCTCGTATATAGACATTTGGGCTGTGCATCAATAAGC 1137
Qy 407 AsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThr 426
Db 1138 TACCATCGTGTAGTAGGGTAGTAAT-----CTTCAACAAATGTATGGAACATAACAA 1191
Qy 427 GlySerProThrThrLysThrIleArgProPheGluSerTyrLysValSerIleValThr 446
Db 1192 AATCTACACAGCACTAGTACCTTT---GATTTTACGAATTATGATATTTTACAAGACTCTA 1248
Qy 447 AspArgGlnSerProProValSerProIleGlnPro-----HisPheIle 461
Db 1249 TCAAGATCGAGTACTCTTGATATTTTACCCTGTTATACGTATATATTTTGGGA 1308
Qy 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
Db 1309 ATCCAGAGAGTTCGAGTTTTCATGGTAAACCAATTGAATAATACCAGAAAGACGTAAAG 1368
Qy 479 TyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArg 498
Db 1369 TAT-----AATCCAGTTTCCAAAGATATATATATATATATATATATATATATATAT 1407
Qy 499 LysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSer 518
Db 1408 GATTCCGAATTAAGAAATTTACCTCCAGAAATTCAGATCAACCAAAATTTATGATCATATAGC 1467
Qy 519 HisIleLeuSerHisPhe---SerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeu 537
Db 1468 CATAGATTATGTCATATACAAAGTATTTCCCGGAGCGGTAACACTACCGGATTA----- 1521
Qy 538 GlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAla 557
Db 1522 -----GTACCTGTATTTCTTGACACATCGAAGTCAGATTTTAAACAATACA 1569
Qy 558 IleSerAspLysIleIleThrMetIleProAlaIleLys---GlyAsnAsnLeuAspThr 576
Db 1570 ATATATTCAGATAAAATCACTCAAAATCCGCGCGTTAAATGTTGGGATAAATTTACCGTTT 1629

```

```
QY 577 AsnSerTysValIleGluGlyProGlyHisThrGlyGlyAsnLeuVal---TyrLeuGln 595
Db 1630 ---GTTCCAGTGGTAAAGGACCGACGACATACAGAGGGGATTTATTACAGTATAATAGA 1686
QY 596 SerGlnGly-----ArgLeuGluIleThrCysGluThrPro 607
Db 1687 AGTACTGGTCTGTAGGAACCTTATTTCTAGCTCGATATGCGCTAGCATTAGAA----- 1740
QY 608 AsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThr 627
Db 1741 ---AAAGCAGGGAATATCGTGAAGACTGAGATATGCTACTGATGCA----- 1785
QY 628 LeuProAsnIleSerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsn 647
Db 1786 -----GATATTGATTGGCATGTAAACGATGCTCAGATTGAGTGCAGGCAAAACATGAAC 1839
QY 648 AsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGln 667
Db 1840 CCAGGT-----GAGGATCTGACATCTAAACCTTTTAAAGTTGCAGAT 1881
QY 668 PheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAla 687
Db 1882 GCTATCACACAGTTAATTTAGCAACAGATAGTTTCGGTTCAGTTAAACATAATTTAGGT 1941
QY 688 AspValSerAsnSer-----IleLeuIleIleAspLysIleGluPheIlePro 703
Db 1942 GAAGACCTTAATCAACATATCTGTTATAGTTTACGTTGACCGAATCGAATTCATCCCA 2001
QY 704 Ile 704
Db 2002 GTA 2004
```

RESULT 15

```
US-11-108-389-73
; Sequence 73, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins with Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 2025
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (mutated)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2025)
US-11-108-389-73
```

Alignment Scores:

Pred. No.:	6.45e-60	Length:	2025
Score:	657.50	Matches:	211
Percent Similarity:	44.67%	Conservative:	120
Best Local Similarity:	28.48%	Mismatches:	293

```
Query Match: 16.99% Indels: 117
DB: 7 Gaps: 31
US-10-783-417-2 (1-735) x US-11-108-389-73 (1-2025)
QY 5 AsnAspAsnAsnGluTyrGluIleIleAspSerHisThrProTyrPheProAsnArg 24
Db 10 AATTAATCAAAATGAATATGAATATATAGATGCGACACCTTCT-----ACTTCT 57
QY 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
Db 58 GTATCCAAATGATTTAAACAGATACCTTTTTCGAATGAGCAACAAATACGCTACAAAT 117
QY 44 ThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
Db 118 ATGATTATTAAGATATATTTAAATAATGTCTCGGGAATGCTAGTGAATACCTCGTTCA 177
QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIleValSer 82
Db 178 CCGAGTACTTGTTCGCGGCAAGAT-----GCAGCTAAGCGCGCAATTGATATAGTA 231
QY 83 GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
Db 232 GGTAAATTAATCTATCAGGTTTAGGG-----GTCCCATTTGTT 267
QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGlyGln 122
Db 268 GGGCCGATAGTAGTCTTTTATCTCACTTATGATATCTGTGGCTTCAGGGGAA--- 324
QY 123 AspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu 142
Db 325 ---AAGAGTCAATGGGAAATTTTATGGAAACAAGTAGAAGAACTCATTAATCAAAAAATA 381
QY 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
Db 382 CGAATATCGAAGCAATAAAGCGCTTTCGGAATTAGAAGGATAGGTAAATTAATACCA 441
QY 163 SerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
Db 442 TTATATCTAACTGGCTTGAAGATGGAAGAA---AATCCATTTTCGAAGTCGAGGTTT 498
QY 183 ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal 202
Db 499 CGAAGTCGAGTGCTTACGAGAT-----GTCCGAAATTCGATTTGAAATCCTCG 546
QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
Db 547 GATAGTTTATTCGCAATATATGCTTCTTTAGAGTGACAAATTTTGAAGTACCATTC 606
QY 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGlu 242
Db 607 CTTACTGTATATGCAATGGCAGCCCACTTCATTTACTGTTTATAAGAGCGCGTCAATT 666
QY 243 LeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
Db 667 TTTGGAGAAGAAATGGGATGG-----TCAACAACACTACT 699
QY 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
Db 700 ATTAATAACTATTATGATCGTCAAAATGAACCTTACTGCAAGATATCTGATCAGTGTGA 759
QY 283 AsnThrTyrArgThrGlyLeuLysAsnLeuAspGluProAsnMetLysTrpSerIle 302
Db 760 AAGTGGTATGAAACTGGTTTAGCAAAATTAAGAGCGCAGCGCTAAACAAATGGGTGAC 819
QY 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db 820 TATAACCAATTCGCTAGAGAAATGACACTGGCGGTTTTAGATGTTGTGCTATTATCCCA 879
QY 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyIleGluValLysGlyIleLys 342
Db 880 AATTATGACACACGACGCTAC-----CCAATGGAAACGAAAGCA----- 918
```

```
QY 343 AsnGluLeuThrArgGluIleTyrThrGluIleAsnPheAspArgLeuProGlnLeu 362
Db 919 ---CAACTACAAAGGAAGTATATACA-----GATCCACTGGCGCGGTA 960
QY 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
Db 961 AACGTGTCCTCAATTGGTTCCTCGTATGACAAAGCACCTTCTTCGGAGGTAGATATCA 1020
QY 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
Db 1021 TCGGTTATTGCGACCAACC---CATGTATTGATTATATATAACGGGACTCACAGTGTATACA 1077
QY 394 GluAsnThrAsnPheGlyAsn---ArgLeuVal-----GlyIleSer 406
Db 1078 CAATCAAGAAGCAATTTCTCCGCTCGCTATATAAGACATTCGGCTGGTCAATCAATAAGC 1137
QY 407 AsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThr 426
Db 1138 TACCATCGTGTAGTAGGGTAGTAAT-----CTTCAACAAGTATAGAACTAATCAA 1191
QY 427 GlySerProThrThrLysThrIleArgProPheGluSerTyrLysValSerIleValThr 446
Db 1192 AATCTACACAGCACTAGTACCTTT---GATTTTACGAATTATGATATTTACAAGACTCTA 1248
QY 447 AspArgGlnSerProProValSerProIleGlnPro-----HisPheIle 461
Db 1249 TCAAGAGGATGCAGTACTCTGTGATATTTGTATACCTCGTGTATACGTATATATTTTGG 1308
QY 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
Db 1309 ATGCCAAGATCGAGTTTTCATGGTAAACCAATTGAAATTAACGAAAGAGCGTTAAAG 1368
QY 479 TyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArg 498
Db 1369 TAT-----AATCCAGTTTCCAAAGATATTTAGCGAGTACAAGA 1407
QY 499 LysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSer 518
Db 1408 GATTCGGAATTAGAAATACCTCCGAAACTTCAGATCAACAAATTTATGATCATATAGC 1467
QY 519 HisIleLeuSerHisPhe---SerLeuPheThrTyrTyrSerTyrValIleGlyLeuGlnLeu 537
Db 1468 CATAGATTATGTCATATAT---AAGTATTTCCGCGCGGTTAACACTACCGGATTA----- 1521
QY 538 GlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAla 557
Db 1522 -----GTACTGTATTTCTTGACACATCGAAGTCGAGATTTAAACAATACA 1569
QY 558 IleSerAspLysIleIleThrMetIleProAlaIleLys---GlyAsnAsnLeuAspThr 576
Db 1570 ATATATTCAGATAAAATCACTCAAAATCCGGCGGTTAAATGTTGGGATATTTACCGTTT 1629
QY 577 AsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuVal---TyrLeuGln 595
Db 1630 ---GTTCCAGTGGTAAAGGACGAGACATACAGGAGGGGATTTATTACAGTATAATAGA 1686
QY 596 SerGlnGly-----ArgLeuGluIleThrCysGluThrPro 607
Db 1687 AGTACTGGTTCGTAGAACCTTATTCTAGCTCGATATGCGCTAGCATTAGAA----- 1740
QY 608 AsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThr 627
Db 1741 ---AAAGCAGGGAATATCGTGTAGACTGAGATATGCTACTGTATGCA----- 1785
QY 628 LeuProAsnIleSerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsn 647
Db 1786 -----GATATTGTATTGCTGTAAGACGATGCTCAGATTCAGATTCGCAAAACAATGAAC 1839
QY 648 AsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGln 667
Db 1840 CCAGGT-----GAGGATCTGACATCTTAAACCTTTTAAAGTTGCAGAT 1881
QY 668 PheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAla 687
```

```
Db 1882 GCTATCAACACAGTTAATTTAGCAACAGATAGTTTCGGTTGCAGTTAAACATAATTTAGGT 1941
QY 688 AspValSerAsnSer-----IleLeuIleIleAspLysIleGluPheIlePro 703
Db 1942 GAAGACCCCTAATTCACACATTTATCTGGTATAGTTTACGTTGACCCGAATCGAATTCATCCA 2001
QY 704 Ile 704
Db 2002 GTA 2004
```

Search completed: December 11, 2005, 23:20:03
Job time : 252 secs